

File ID	uvapub:66318
Filename	Thesis
Version	unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	PhD thesis
Title	Legal theory, sources of law and the semantic web
Author(s)	A. Boer
Faculty	FdR: Leibniz Center for Law, FdR
Year	2009

FULL BIBLIOGRAPHIC DETAILS:

<http://hdl.handle.net/11245/1.333378>

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content licence (like Creative Commons).

195

LEGAL THEORY, SOURCES OF LAW AND
THE SEMANTIC WEBITION

A. Boer

Frontiers
in
Artificial
Intelligence
and
Applications

LEGAL THEORY, SOURCES OF LAW AND THE SEMANTIC WEB

Alexander Boer



ISBN 978-1-60750-003-2
ISSN 0922-6389

IOS
Press

IOS
Press

Legal Theory, Sources of Law & the Semantic Web

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom

ten overstaan van een door het college van promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel der Universiteit
op donderdag 25 juni 2009, te 14:00 uur

door Alexander Willem Frank Boer
geboren te Alkmaar

Promotiecommissie

Promotor: Prof. dr. T. M. van Engers

Copromotor: Prof. dr. J. A. P. J. Breuker
Dr. R. G. F. Winkels

Overige leden: Prof. dr. T. J. M. Bench-Capon
Dr. B. Bredeweg
Prof. dr. mr. H. Prakken
Prof. mr. P. P. Rijpkema
Prof. dr. G. Sartor
Prof. dr. B.J. Wielinga

Faculteit: Rechtsgeleerdheid

Legal Theory, Sources of Law & the Semantic Web

Alexander Boer

*Leibniz Center for Law
University of Amsterdam*

IOS Press

This is a special electronic edition made available to the University of Amsterdam in compliance with article 27a of promotion regulations. The commercial edition is published as volume 195 of *Frontiers in Artificial Intelligence and Applications* by:

IOS Press
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
e-mail: order@iospress.nl

ISBN 978-1-60750-003-2
Library of Congress Control Number 2009924846
doi:10.3233/978-1-60750-003-2-i



SIKS Dissertation Series No. 2009-11

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

Contents

Chapter 1. Introduction	1
1.1 Motivation of this Book	2
1.2 Sources of Law and the Semantic Web	3
1.3 Research Approach	6
1.4 Outline of this Book	10
Chapter 2. Knowledge Engineering and the Law	13
2.1 Introduction	14
2.2 Knowledge Representation	17
2.3 Ontology	32
2.4 Epistemology	40
2.5 Argumentation and Information	48
2.6 Conclusions	53
Chapter 3. Knowledge Components	57
3.1 Introduction	58
3.2 The Semantic Web	60
3.3 Knowledge Components and Revision	64
3.4 Description Logic and Interaction	67
3.5 About Relevant Knowledge Sources	79
3.6 Conclusions	84
Chapter 4. Agency and the Law	87
4.1 Introduction	88
4.2 Institutions and Rules	92
4.3 Intention and Action	96
4.4 Law as an Institution	100
4.5 Norms, Normality and Normativity	108
4.6 Legal Normative Rules	115
4.7 Agents and Action	123
4.8 Interaction Between Agents	134
4.9 Conclusions	142

Chapter 5. Representation of Sources of Law	145
5.1 Introduction	146
5.2 Sources of Law and Legal Rules	148
5.3 Sources of Law and the Semantic Web	164
5.4 Summary and Conclusions	185
Chapter 6. The Normative Order	193
6.1 Introduction	194
6.2 Representation of Normative Rules	198
6.3 Normative Conflict	203
6.4 Contrary to Duty Obligations	210
6.5 Betterness in OWL DL	214
6.6 Choice between Legal Rules	226
6.7 Elaborating on the Intended Normative Order	235
6.8 Conclusions	256
Chapter 7. Summary and Conclusions	261
7.1 Introduction	262
7.2 Summary	262
7.3 Conclusions	269
7.4 Research Questions	275
Chapter A. About MetaLex and the LKIF Ontology	279
Chapter B. Rechtstheorie, rechtsbronnen en het <i>Semantic Web</i>	283
B.1 Inleiding	283
B.2 Kennisrepresentatie	284
B.3 Het recht als onderwerp	285
B.4 Rechtsbronnen en hun inhoud	287
B.5 Redeneren over de bedoelde normatieve orde	288
Bibliography	291

Chapter 1

Introduction

1.1. Motivation of this Book

Any system simple enough to be understandable will not be complicated enough to behave intelligently; and any system complicated enough to behave intelligently will not be simple enough to understand. (George Dyson, 2008)

Driven by the Semantic Web, increasing legal convergence, and an increasing pace of organizational change in public administration, the interest in legal knowledge representation in public administrations is gradually increasing but also changing in nature.

Initially interest was focused on the utility of the computer systems built using a knowledge engineering approach; More recently the focus shifted to the potential utility of knowledge representation for comparative and maintenance purposes, and for increasing the efficiency of the organizational change process itself. This change in focus has consequences for the knowledge engineering community.

The concept of *knowledge* representation is strongly linked to the ability of computer systems to make decisions. The basic theses of knowledge representation can be summarized as follows:

1. it is possible to implement computer systems that are capable of making decisions that would require relevant knowledge – in this case of the law – if made by a human;
2. therefore the computer system in some way contains knowledge, in the sense that we *ascribe* knowledge to it: it exists as an agent on the knowledge level (cf. [209]);
3. if this (ascribed) *knowledge base* is made explicit and structured in the right way, then it consists of coherent knowledge components, at least some of which are reusable by different decision-making systems that require the same knowledge (cf. [120]).

The focal point of the Semantic Web vision is however not on improving on the state of the art in automated decision making, but on the efficiency of maintenance and the possibility of sharing and comparing knowledge components. The problem the knowledge engineering community faces is the problem of representing this knowledge in *the right way* in order to achieve reusability outside the original context of use. This is essentially an exercise into making unmanageable complex problems manageable by decomposing them into less complex problems.

Compared to the standards set by knowledge engineering research, fielded systems in public administration and elsewhere that use explicit knowledge representation to support decision making processes are technically and theoretically straightforward. The required transparency, and the challenges real world knowledge representation pose for the people implementing them, act as a natural limit to the complexity of these systems. Since the combination of knowledge of knowledge engineering and law is a rarity in the available work force, the complexities of legal theory and the interpretation of sources of law – like legislation, official directives, and case law – are largely ignored.

This book is an attempt to construct an integrated conceptual framework for the

application-neutral and problem-neutral representation of sources of law using Semantic Web technology and concepts, and some technically straightforward extensions to Semantic Web technology based on established practices found in fielded applications. To construct this framework, I disentangle some problems that are in my view too often mixed up in legal theory and – in extension – legal knowledge representation.

The purpose of this framework is to provide a theoretical background for the creation of reusable and maintainable knowledge components representing knowledge of sources of law on the Semantic Web. These components should form a basis for the development for computer applications supporting straightforward, routine decision making problems using traditional methods.

The next section sketches relevant recent developments, in the field of law and public administration, and the objectives of some projects I have been involved in over the last decade. This section functions as a background against which the relevance of the research themes addressed in this book can be understood.

Section 1.3 introduces the research approach and research themes addressed in this book, and section 1.4 gives an overview of the book itself and the topics addressed in it.

1.2. Sources of Law and the Semantic Web

The increasing ubiquity of the Internet in society has put a pressure on public administrations to make previously paper-based services and sources of information available on the Internet. If you can do your bank business online, why should you expect less from your government?

The migration of public administration to the Internet is loosely referred to as *eGovernment*, or electronic government. The enabling technology spurring recent interest in electronic government is the *Semantic Web*, a vision of a next-generation World Wide Web – complemented with certain design principles and core technologies – focused on loosely connected *services* and *knowledge components* instead of pages.

The services in question are in an administrative setting generally public legal acts, performed by public legal personalities, based in formal legislation. Legislation defines what the core functions of public organizations are, and what services it provides. It guides how the organization subdivides itself into administrative units, how it organizes business processes inside the organization, and eventually how the functions of the organization are realized by civil servants and computer systems. Everything the public administration does that cannot be done by a private organization or person is normally speaking based in legislation.

In many cases eGovernment is an internet-based portal connected to old, slow, mainly paper-based decision making processes. But electronic government cannot be just a matter of mere window-dressing. Inside public administrations, and on the interfaces between them, ICT and Internet also have a large impact. Some decision making processes are nowadays assisted by computer applications, and others are more or less autonomously performed by the computer. Many tax declarations – sometimes more than 80% of them – are for instance never handled by a human being. The computer decides, randomly or based on properties of the declaration that suggest

intentional or unintentional misrepresentation, which declarations should be looked at by a civil servant.

Business process design and design of specialized computer systems are both usually based on explicit *models* in modeling languages, like the Unified Modeling Language (UML) or, more recently in the context of the Semantic Web, the Web Ontology Language (OWL), of what the business process or computer application should achieve. These models are supposedly used as a specification of the objectives of an organizational change process or application development process. When legislation changes, these models are updated, and the organization's structures and computer programs have to be changed to conform to the models.

In the past these changes were conceived of as temporary interruptions of long periods of everything staying the same. This was certainly the case when the adaptation of existing computer systems was still considered a frightening prospect: things did change but the changes were carefully orchestrated to not impact existing procedures and computer systems. But as the perceived capacity of organizations to organize change processes increases, and the number of fielded computer applications increases, so does the pace of change in legislation directly affecting existing computer applications.

Tax legislation is for instance changed every year, leading to continuous adaptation of relevant computer applications for next year and the years after that, while the legislation of the present year and previous years is still being applied. In the business process design literature, awareness of this phenomenon has led to a new conception of the organization as an entity that is constantly in the process of changing: the organization is constantly conceptualizing and comparing what it is and what it is becoming.

Attention for *knowledge representation* of sources of law – the subject of this book – is very often triggered by such administrative change processes driven by new legislation and other sources of law, such as case law and internal written policies. Knowledge representation is seen as a means to potentially reduce costs and increase efficiency through increased control over the knowledge dimension of the change process.

The increased focus on change is however not just driven by technological possibilities and pipe dreams, and the perception that the capacity to change has increased. The change processes triggered by the legal system are increasingly expensive, especially if they involve changes in ICT infrastructure.

The process of drafting consistent, coherent, and effective legislation itself is indeed getting more complicated, as is that of efficiently using, enforcing and applying valid law. The very fact that legislators recognize each other's activities as legitimate, embrace *legal pluralism*, and want to cooperate and compete with each other, tends to complicate the operations of law.

The increasing legal convergence between governments in the European Union, and the increasing traffic of people over borders of jurisdictions, inevitably leads to an increased interest in the problem of knowing, comparing and harmonizing legislation.

Administrations need to know and understand legislation of friendly governments to be able to assist citizens and reduce negative impact of permitted movements between jurisdictions.

Employees of tax administrations in the EU are for instance increasingly confronted with requests that require them to understand European regulations and directives, and regulations of other EU member states; They need to react to the consequences of increased movement of people, products, and money between EU member states and increased harmonization between tax authorities in Europe.

Subject to certain conditions and exceptions, value added tax for instance has to be paid only once in the member state of choice of the owner of the goods, according to a EU directive. The directive has as a costly administrative side-effect that tax administrations employees suddenly need to be able to apply the VAT rules of all 25 member states to know whether the conditions have been met or the exceptions apply.

Today's public administrations do not yet have an effective answer to demands such as these. And neither does the market.

The market is solving its own comparison problems. Global companies offer products and services in many jurisdictions at the same time, and the product or service has to meet the provisions of all jurisdictions in which it is offered.

The issue is however not just one of compliance. Different regulations also lead to large differences in competitiveness for the product or service. For a financial product, for instance, it is considered important to qualify for tax deductions that make the product more attractive. For a medical product it is for instance important to know whether it will be possible to sell it over the counter without a prescription. For computer applications it is for instance important to know to what extent it will be protected by copyright, and which license conditions are legally enforceable. For outsourced administrative services it is for instance sometimes imperative that privacy of client data is legally protected to the same extent as it would be in the jurisdiction where the client data comes from, and this limits from and to which countries services can be outsourced.

This increased attention for comparing legal systems is for instance evidenced by the number of consultancy firms that advertise their knowledge of multiple legal systems to companies on the World Wide Web. In addition, there are some initiatives – often initiated outside the Computer Science and Law community – for constructing international legal *ontologies* that expose subsumption relations between legal vocabulary in multiple jurisdictions, and make law comparable, on the Semantic Web.

For public administrations the idea of the Semantic Web has great potential. Public administrations that provide different services based on the same legislation can share the effort of making a knowledge representation of that law, or, even better, delegate that responsibility to the market or even to the legislator: future legislation may come with an embedded knowledge representation, and of course over the Internet, to its institutional users.

In addition to that, if you can use knowledge representation for, let's say, calculating how much a taxpayer owes the government in income taxes, and the legislator makes available a new version of the relevant knowledge representation, then it would obviously be possible to compare the old and the new version to see what has changed. Even more ambitiously, if a third country sends a knowledge representation of its solution to the problem of taxing income, then the administration would appear to be able compare it with its own knowledge representation, and more or less automatically

deduce what relevant differences there are.

My recent work for the Dutch Tax and Customs Administration (DTCA; cf. [42]) was for instance clearly related to the huge *change process* triggered by the complete overhaul of the Dutch income tax law in 2001. DTCA knowledge representations intended for use in decision support systems were expected to specify – and justify – how business processes were grounded in legislation, and provided input for the drafting process through formal verification and simulation of draft legislation. The *Juridisch Loket* (cf. [272]) project on *pro bono* legal assistance, and the *DURP* project on spatial planning (cf. [40]) were also driven by an overhaul of legislation.

An ambitious quest – one we tried to realize for the DTCA – is to integrate the process of legislative drafting, including evaluation of formal and executable models of legislation, and deployment of these formal models as software for the Dutch Tax and Customs Administration, into a single architectural framework. Although the project was successful – at least by the standards set for such projects – there is still a lot to do.

Some of the recent knowledge representation projects I was involved in (E-POWER, Estrella) included the *comparison* of representations of similar pieces of legislation from different jurisdictions. Although pairing up different decision support systems solving the “same” problem is certainly a way to gain insight into relevant differences, the field of legal knowledge engineering really has no theoretical framework for comparison of knowledge representation for the purpose of exposing relevant differences in the law.

The idea that this is possible and meaningful, presumes a lot of the degree of inter-coder reliability that can be achieved by knowledge engineers: if there is *no* difference in the source of law represented by two different knowledge representations from two different authors, do we really expect no difference in these knowledge representations? This calls for strong methodological commitment and standardization of approach in legal knowledge representation, which does not exist in practice.

1.3. Research Approach

The Semantic Web, and the assumptions behind it, demand a lot of the structure of knowledge bases.

The Semantic Web is no longer about the knowledge base, as the driving force behind a computer application that is capable of making decisions that require knowledge, but about the flexible interaction between knowledge components made available and maintained by different suppliers. Applications will no longer contain all knowledge they use, but instead they will download the latest version of knowledge components from various locations on the Internet on demand. This is the idea of *federalized* knowledge bases, or of *knowledge syndication* when considered as a business model.

This book is a reflection on the representation of knowledge about the content of sources of law. The following assumptions are taken for granted in this work:

1. *legal* knowledge is mainly based on the formal sources of law, although there is also *legally relevant* knowledge, and;

2. the sources of law are the main locus of legal change, and therefore the appropriate locus of maintenance activity.

Certain design principles based on these assumptions have to be accepted as a given. Reusable components aim to represent a specific source of law, and the content of the component can be traced to and justified by the source of law – this is *isomorphism* as it is used in Computer Science and Law. Maintaining isomorphism to a high degree is one of the tools we have as knowledge engineers to increase methodological rigor and make maintenance manageable in knowledge representation.

An appropriate selection of these components, plus additional relevant knowledge from other sources, form the knowledge base of a decision support system. A new consolidated version of the source of law corresponds with an updated knowledge component that replaces the previous one, within the existing structure of a knowledge base, while minimizing impact on the other knowledge components present in the knowledge base.

This is a more or less traditional setup, which will be treated as a given – as part of the problem – in this book. This is the role legal knowledge engineers can successfully adopt in the Semantic Web: represent the sources of law, not the whole of legal problem solving, and one meets the expectations and needs of the market and respects the implicit division of labor presumed by the Semantic Web vision.

This book is a reflection – not a *report* – on ten years of applied research in legal knowledge engineering. Throughout the last ten years I have formulated answers, in the form of a conceptual framework, to three related questions relevant to the Semantic Web idea as applied to sources of law:

1. If a reusable knowledge component is to be used in different settings, and to help realize different problem-solving competences, what part of a knowledge representation is the reusable part and what part is specific to the setting or problem to be solved? This is an inquiry into what is *essential* in knowledge of the sources of law, based on the assumption that this quality determines the degree of reusability of knowledge.
2. How does one integrate existing design principles, core technologies, and technical standards of the Semantic Web – such as its naming and addressing standards, based on the uniform resource identifier or URI, and the description logic OWL DL – with legal knowledge representation?
3. What knowledge should be considered specifically legal, and what knowledge should be considered as a given? This is in essence an inquiry into the appropriate division of labour: the notion of *federalization* of knowledge representation presupposes that reusable knowledge components should be made available by parties specifically competent to do so;

This book adopts the strict distinction between *ontology* – about the *object* of our knowledge – and *epistemology* – about the processes by which knowledge is attained – maintained by for instance by Breuker i.a. in [67]. The distinction between these two aspects of knowledge will be worked out throughout chapter 2.

The appropriate object of study when reflecting on knowledge components for the Semantic Web is *ontology*. Ontology cannot be completely disconnected from

epistemology – and this book will address practical reasoning approaches and problems where relevant – but one can decide to be extremely cautious when making ontological commitments. The motivation for making ontological commitments should not be to achieve a certain level of problem solving competence. The pragmatics of problem solving should be kept at arm’s length.

The conceptual framework sketched in this book will therefore leave noticeable gaps if one approaches it as a grand, unified theory of legal reasoning: these are intentional, and filling them in is left as an exercise for each individual reader.

Many individual solutions chosen in this book have a pedigree in literature, and certainly do not aim to improve on the state of the art in automated legal reasoning. As a conceptual and methodological framework, and an attempt to make legislation Semantic Webbable, this work is however new.

Law must be the knowledge domain *par excellence* for testing the concepts of the Semantic Web. Legislation, with its internal references, import mechanisms, applicability conditions, and intricate version management solutions, is the closest thing to a tested and tried, paper-based semantic web in existence.

The Semantic Web is however not designed with law and legal knowledge representation in mind, which brings us to the second question.

The Semantic Web technology of central importance for knowledge representation is OWL DL, a reasonably expressive subset of the Web Ontology Language that can be interpreted as a first order monotonic *description logic* with reasonably attractive computational properties. OWL DL does not support propositions about propositions, for sound engineering reasons.

The monotonicity of OWL DL is attractive from an engineering point of view: it permits problem decomposition in knowledge representation. Monotonicity is essential for the idea of composing knowledge bases from separate knowledge components, in analogy with the design of physical artifacts from physical components with a transparent and predictable interface.

Classical monotonic logics are however not very plausible models of real world decision making. Sometimes we can prove something once and for all; Most of the time we can only construct falsifiable arguments. If we construct these arguments from formal rules, then these rules must be *defeasible*. This is true as much in law as it is true in any other domain of enquiry, but in legal practice, where reflection on the applicability of rules is common, it is especially obvious.

Moreover, the sources of law even explicitly address the falsifiability of legal argument, and occasionally formulate constraints on reasoning that explicitly presume the possibility of falsification of what would otherwise be acceptable as a valid argument. The formulation of these constraints often involves propositions about propositions.

Even if one decides *not* to solve the problem of defeasible reasoning, and accepts that there are many more or less valid ways in which one can – defeasibly – interpret and use the same source of law, perhaps even the same knowledge representation of it, one should at least try to represent these constraints in order to maintain isomorphism of knowledge representation and source of law.

The third question – the distinction between legal and non-legal knowledge – can be approached both from the point of view of *ontology* as of *epistemology*.

The *epistemological* dimension of this question is the more important one. A basic assumption about epistemology underlying the conceptual framework sketched in this book is perhaps a provocative one: *legal reasoning* as such – as a special kind of mental process – does not exist, and we do not have to account for a special legal epistemological theory.

Everybody understands that representation of the traffic rules is not enough to design an intelligent system that can drive a car; There is no reason to expect that law does give rise to a *legal logic* that accounts for legal reasoning. Legal reasoning is straightforward common sense reasoning, but with legal knowledge.

On the other hand the law instructs us not only on what to do but also often on how to reason. But law only restricts the degrees of freedom we have in reasoning, through explicitly formulated rules; We do not have to interpret it in terms of a theory of legal reasoning.

If the Semantic Web does not accommodate defeasible reasoning, one shouldn't have to introduce it to accurately represent knowledge of the law. Falsifiability is not unique to law. If OWL DL is good for anything else, it is also good enough for law. The same applies to higher-order reasoning involving propositions about propositions: OWL DL does not accommodate it, but reflection on propositions is not unique to law. The objective of this book is to represent legal knowledge in OWL DL, without extending or changing the semantics of OWL DL.

Constitutiveness and applicability are for instance familiar concepts in legal knowledge representation, but are normally conceived of as higher-order observations about propositions. In this book they are treated as first order properties of objects: legal acts and other legal *things* take logical priority over legal *facts*, which are left implicit. *Legal rules* are indeed reified propositions if one wishes to see them in that way, but they are never interpreted as propositions in reasoning.

The question is how one can preserve a reasonable degree of descriptive fidelity in knowledge representation without these devices. Special legal logic is unnecessary for most fielded applications, which rely heavily on context of use and a very narrowly defined problem solving competence to avoid complex reasoning techniques.

There is of course also an *ontological* dimension to the distinction between legal and non-legal knowledge: for this purpose the concept of the law as a set of institutions is introduced. The source of law uses legal terms that belong to the institution, and *legally relevant* ones that are on the interface between the institution and the outside world. The institution is distinguished from the normative order it intends to create. Normative order is the ultimate function of law – and an essential one – rather than a part of it. Important parts of the normative order are found only *between the lines* in the sources of law.

The first and central question – how to design knowledge components representing sources of law for *reusability* – is of course not one which can be definitively answered. There are a number of known threats to reusability, and we can do our best to contain them.

I already addressed the problem of keeping the pragmatics of problem solving at arm's length. There are certain relevant contextual assumptions that affect the way in which the knowledge engineer represents knowledge for a specific problem

setting. An important objective of this book is to make a survey of relevant concepts in legal theory with the approach chosen for this book in mind, in order to remove – often epistemological – aspects of these concepts that are non-essential for the mere *representation* of the sources of law in the form of rules.

The distinction between the legal and non-legal made by the second question addresses another reusability concern: if legal knowledge engineering invents its own solutions for every relevant aspect of legal reasoning, or any kind of relevant domain knowledge, it is likely to end up with products that are hard to fit in as a component in more general knowledge engineering projects. Law is more often than not only *part* of the problem context. Respecting existing, more general approaches, even if they are hard to use, should be an objective in itself for knowledge engineers.

A trivial reusability concern, already addressed by the first question, is a lack of technical interoperability. Semantic Web standards make technical interoperability possible, but at the same time impose rather arbitrary restrictions on the expression of knowledge. Remaining within the limits set by the standards should be an objective in itself.

Most pressing is however giving falsifiability in general a place without declaring everything potentially falsifiable. Unchecked falsifiability is a license for bad engineering, or for declaring reuse of knowledge in principle impossible. As will be demonstrated in this book, defeasibility is not only a problem because OWL DL does not allow it. Defeasibility is also a central issue for the distinction between ontological knowledge and other, contingent knowledge, and even the distinction between legal and legally relevant knowledge can be understood in terms of defeasibility.

1.4. Outline of this Book

Chapter 2 of this work examines the field of knowledge engineering. It starts gently, for readers not familiar with the field, with an introduction into knowledge representation and knowledge-based systems. The chapter also sets the stage for the rest of the book: it introduces ontology, epistemology, ampliative or abductive reasoning, ontological stratification, and abstraction as used in this book.

Chapter 3 presents the idea of a knowledge component on the Semantic Web, based on OWL DL extended with some reasoning mechanism for autoepistemic reasoning. It defines the concept of a source of law, the object of the knowledge to be represented, and introduces certain notational conventions and an important distinction between several different types of logical rule used throughout the book.

Chapter 4 gives an account of certain fundamental theoretical concepts and issues in law and legal knowledge engineering. It introduces institutions and their rules, normality, and normativity, and explains some relations between the rules of the institution, intention, and action. The primary purpose of this account is to separate the rules themselves from assumptions about their use in planning. For general knowledge engineers it serves as an introduction into the theoretical problems of this field. For legal theorists most concepts will be well known, but the chapter will give a new perspective on some issues.

In this work the device of *ontological stratification*, introduced in chapter 2, is used to structure legal knowledge, inspired by the use of institutions and constitutiveness in legal theory. Stratification also plays a key role in containing defeasibility in reasoning: the only inferences that are permitted to be defeasible are those that bridge ontological strata. The central stratum of chapter 4 is the legal institutional reality.

Law is a domain of *institutions* whose main purpose is to create an intended normative order by formalizing it. The legislator as engineer of his own institutional reality is responsible for the rules and structures of the institution being coherent and transparent, understandable, justifiable, and capable of effectively and efficiently achieving the intended normative order. The legislator shares the problems of the legal knowledge engineer: both manage an intangible and very complex artifact whose structure must be designed for understandability, transparency, reusability, maintenance requirements, etc.

A central ontological theme of this book is the representation of the institution, and the relation between the institution and the formal sources of law that express it. What legislative action does, is to change the rules and structures of the institution. What the legislator *intends* is however to change prevailing normative order, towards an intended one, as chapter 4 explains.

Chapter 5 focuses on the logical representation of information about legal rules and the problems legislators encounter in managing a large body of legal rules, the solutions they have found for these problems, and the consequences this has for knowledge engineers who try to represent the meaning of these rules in logical form. The chapter is the main body of the work from a knowledge representation point of view.

In this chapter the *legal rule* is dissociated from the fragment of a source of law that represents it and from both the logical rule of chapter 3 and the rules of the institution in 4. It also refines the relationship between the logical rules introduced in chapter 3 and the rules of the institution.

The appropriate relation between the source of law *qua expression*, the legal rules *qua institutional entity*, the intended effect of those legal rules on the normative order, and the logical sentences of the knowledge representation is a central theme in this work. My work on MetaLex, also reported in chapter 5, revolves around the sometimes non-trivial relation between changes to the sources of law and changes to the rules and structures of the institution.

Chapter 5 concludes with a reflection on the relationship between sources of law and knowledge components composed of logical rules.

Chapter 6 discusses the normative order the legislator intends to create through the institution and its rules. In this chapter the most familiar ingredient of normative order created by law – the normative rule – is worked out in OWL DL. The chapter also argues that it is necessary to ascribe an intended normative order to the actions of the legislator to explain non-trivial varieties of legal reasoning such as court adjudication and comparative law.

The usefulness of decomposition of systems into reusable components largely depends on the recognition of the functions of the individual component. These functions of the source of law can be inferred from the intended normative order, in the understanding that ascribing such intentions to the legislator is not a clear-cut problem. The

representation of normative order, beyond the fragment made explicit by normative rules, still presents us with a number of open research questions; Chapter 7 presents some suggestions for further research into the representation of normative order, besides its obvious purpose of recapitulating the essentials of the preceding chapters.

Throughout this book regular references are made to two Semantic Web-related public specifications that I have been closely involved in: the MetaLex CEN/ISSS standard and the Legal Knowledge Interchange Format (LKIF) ontology. Appendix A positions this work in relation to MetaLex and the LKIF ontology.

Chapter 2

Knowledge Engineering and the Law

2.1. Introduction

Knowledge Engineering is the technique applied by knowledge engineers to build *knowledge based systems* (KBS). Given a knowledge-intensive task normally performed by knowledgeable human problem solvers, the knowledge engineer attempts to model the domain knowledge and problem solving techniques of the human problem solver, and uses this model to implement a KBS capable of performing the knowledge-intensive task.

The KBS in principle replaces the human problem solver capable of performing the task by himself, either completely or – more commonly – by allowing less skilled workers to perform the same task assisted by the KBS. The introduction of a KBS is typically intended to take boring routine tasks out of the hands of scarce experts so that they can spend more time on the hard problems. If KBS are for instance capable of handling the great bulk of tax applications without any human intervention – a stage that has already been reached in a number of countries – the experts in the tax administration will get to spend more time on complicated and suspect cases.

Scarcity of expertise seems inescapable in the legal field in the foreseeable future: any increase in efficiency of the legal system will presumably immediately be canceled by increasing demand. The argument for this observation is simple: if the legal system is the expensive *last resort* solution for maintaining normative order in a society, as is often argued, any reduction in the costs of this solution generates increased demand. If the fielded KBS is cheaper than the people it replaces it is therefore reasonable to expect that the use of KBS increases the quality of jobs in the legal field, while making the legal solutions more accessible.

Effectiveness is however usually the more important reason to field KBS in the legal system: *uniformity* (or *stare decisis*) in legal decision making is a goal in itself in the field of law, and uniformity brings down the overall costs of the legal system by giving people less incentive to reject or appeal against its decisions.

The legal expert's role gradually shifts from decision making to setting uniform decision making policies for the KBS, i.e. to producing (decision making) knowledge instead of decisions. This shift is significant, and is mirrored by a recent shift in attention in knowledge engineering and knowledge management literature (cf. [191]): the shift from attention for the problem solving competence of KBS to attention for the decentralized maintenance of reusable knowledge components.

In this chapter the discipline of knowledge engineering, and its subject matter, is introduced, to help the reader position the observations and judgments in the following chapters in their context of use.

The framework sketched in this chapter is however rather specific for a number of goals. Reasoning and knowledge as conceived in this book are based in the concept of reasoning as *problem solving*, and knowledge as a resource in problem solving. Section 2.2 introduces this view of knowledge, and introduces knowledge components.

Since the quest for reusability puts the question of falsifiability of knowledge on the center stage, the distinction – made in section 2.2.5 – between *ontological* knowledge that can be considered not falsifiable, and other knowledge that is considered falsifiable, receives a lot of attention. The concepts of *ontology*, *abstraction* and *ontological*

stratification play a key role in this, and are explained in section 2.3.

The concept of problem solving, and the role of falsifiable inference in problem solving, is addressed in section 2.4 about epistemology. In addition a distinction is drawn between knowledge and reasoning, the subject of this book, on the one hand, and information and argumentation on the other hand in section 2.5. This distinction is intended for scoping purposes: although there is a lot one should know about information and argumentation to design good LKBS, this subject the beyond the scope of this book.

2.1.1. Legal Knowledge Based Systems

A *legal knowledge based system* (LKBS) is a KBS that mainly deals with legal subject matter.

Legal knowledge based systems are mostly used for routine tasks, usually by public bodies specialized in a limited number of decision processes constrained by law relating to one topic. I have for instance worked in applied research projects in taxation (e.g. [270, 51, 50]), social security (e.g. [272, 271, 287]), and court (e.g. [60, 64]) settings.

Unfortunately no empirical studies along the lines of [124] have been undertaken specifically on *legal* KBS: there appears to be no hard data about their prevalence, or about how they fare over time. Rare empirical research into the quality of decisions made by LKBS suggests that matching or exceeding human performance is quite feasible (cf. generally [129]). Generalizing from the results of [124] one might speculate that many more than the ones discussed in academic literature qualify as LKBS¹, and that two out of three systems fall out of use within 5 years. In the field of law a system may be discarded after some years because the relevant sources of law have changed so considerably that adapting the LKBS becomes too costly or time-consuming.

LKBS in combination with *Semantic Web* technologies, and information standardization in general, are, under the moniker *eGovernment*, often proposed in political-administrative circles as *potentially* the most effective non-legal solution for reducing the *administrative burden*, the costs of carrying out administrative activities that one would not carry out in the absence of regulation (cf. the bibliography in [7], or the European Commission's view in COM2007/23, p.12, which mentions "intelligent portals").

This administrative burden is often conceived of as a form of economic *deadweight loss*. Note that this deadweight loss is incurred both through taxes spent on administrative processes within the government and through direct administrative costs made by citizens and businesses. Political attention appears to be directed towards the direct costs of businesses: as pointed out in [7] reduction of the administrative burden may simply move the deadweight from businesses to the government instead of reducing it. The justification of this belief in reducing economic deadweight loss through "intelligent portals" is beyond the scope of this book.

While the impact of LKBS on institutional effectiveness and efficiency is hard to judge, in part because of the hidden demand issue noted earlier, LKBS *could* have a

¹An even more daring speculation: there is probably also no positive relation between output of scientific publications about a system and its real world success.

disruptive, even revolutionary, effect on an inter-institutional level.

IT in general turns out to be a major force in leveling playing fields between governments competing for favour in some areas: we for instance regularly see very unlikely candidates scoring high on *eGovernment* rankings. In [279] we see for instance Equatorial Guinea, Liberia, Sierra Leone, Bhutan, Eritrea, Gabon, and Ethiopia outranking not just wealthier, but also clearly IT-savvy nations such as Japan, France, Sweden, Denmark, and Norway in 2007. The most obvious explanation for these outliers is that these countries invest more (relatively speaking) in eGovernment for competitive reasons.

Competitive use of eGovernment more often than not involves automated decision making processes. Electronic tax filing appears to be a major contributing factor to increased tax competition between countries, as it becomes increasingly easier for companies to declare taxes and duties wherever it is judged to be most advantageous. The LKBS helps to advertise the legal services a country provides, by making them easy and efficient to use. LKBS that rearrange business operations continually to optimize taxes – declared to government administration LKBS – are foreseeable. In general LKBS could make traffic of persons, goods and services over national borders a lot easier, provided that systems can interface with each other, and their inputs and outputs can be easily compared. But this turns out to be a major problem, for reasons described in section 2.2.4.

In the many areas where governments however do not compete with each other for the customer's favour, eGovernment easily becomes merely window dressing instead of a new window: an internet-based portal connected to the same old, slow decision making processes (cf. generally [71, 8, 7]).

The field of legal knowledge engineering specializes in knowledge engineering topics specific or typical to the development of LKBS. Some of these topics are:

Formal sources of law play a central role as written *knowledge sources* in LKBS development and maintenance. In law the sources telling us how to think and how to act are *explicit*. This characteristic seems to make law eminently suitable to KBS, and law is indeed one of the oldest areas of application of KBS (viz. [200]). At the same time, legal constraints on “how to think” do not necessarily make building a system, or formulating generic theories about the design of such systems, *easier*. So-called *tacit knowledge* (cf. [68, 191]) – which is hard to make explicit – of experts plays a larger role in other disciplines. In most KBS projects the required knowledge has to be “discovered” and “acquired”: in law it is more readily available (cf. section 3.5.1).

In law, it is also often change in the sources of law, sometimes just a court verdict contrary to the interpretation of the LKBS, that dictates an update of the LKBS. The law continually makes the knowledge it creates explicit in the form of (changes to) the sources of law, and itself sets the deadlines for when it should be used in decision making. This is uncommon in many other application domains, where the represented domain usually must remain constant for fairly long times, or only changes at the instigation of the KBS user², to make the deployment of a KBS

²The introduction of new components in the product catalogue for instance prompts the intro-

justifiable.

The distinction between prescriptive and descriptive is important in all types of reasoning, but the field of law has turned the specification and categorization of prescriptive knowledge into an art form. Most reasoning tasks require great amounts of invariant, descriptive domain knowledge, and a few goals, constraints, requirements, norms to prescribe what the generic solution looks like. Law is about mostly concerned with the organization of prescriptive knowledge, and borrows descriptive knowledge from other knowledge domains when needed.

The central role of justification is an essential characteristic of law. The procedurally and substantively correct method is more important than the quality of the outcome, for which no independent standard is available. Most academic disciplines apply scientific method to some domain of interest and are primarily interested in correct outcomes, evaluated against some external standard of correctness, regardless of the method followed. The outcome without a proper justification in terms of the prescribed method is of little value in law.

The distinction between facts and claims has led to a wide array of logics, frameworks, and theories dealing with *adversarial* settings where different parties disagree on the facts. Many of these efforts were initiated by people with experience in Computer Science & Law (cf. [236]). A medical KBS will for instance typically never even distinguish between facts and the claims made by the user(s) of the system, and is not built to handle logical conflicts arising from contradictory claims relating to the same facts. We might refer to this as the presumption of *cooperativeness*. In LKBS application areas questionable and contradictory claims about the facts are commonplace.

A legal knowledge representation makes well-motivated representation choices with respect to these topics. In the next section we focus on the nature of these representation choices.

2.2. Knowledge Representation

Legal knowledge engineering aims to express legal knowledge in a language that can be used for communicating and storing what we know, and for making knowledge productive in the form of LKBS. If we express knowledge in such a language we call the activity *knowledge representation*. If a language was designed for representing knowledge we call that language a knowledge representation language. The result of the knowledge representation activity is a knowledge representation or *knowledge base*, depending on context.

Knowledge representation languages can be considered as a special kind of language specifically designed to express knowledge. Some knowledge representations are so frequently used that almost no one recognizes them as being a knowledge representation. Musicians for instance use a knowledge representation language that is centuries old and allows musicians to express and exchange ideas. Mathematicians, physicians, chemists, pharmacologists and others have also introduced a plethora of knowledge

duction of new component descriptions in a parametric design system

representations. The reader will recognize this formula from his or her high school period:

$$F = \frac{\delta p}{\delta t}$$

Momentum p measures the motion content of an object, and is the product of an object's mass and velocity. Momentum doubles, for example, when velocity doubles. Momentum and force are related by the fact that force F is the rate at which momentum changes with respect to time. If m is constant – which is often the case in application scenarios – then the following holds:

$$F = \frac{\delta p}{\delta t} = m \frac{\delta v}{\delta t} = ma$$

The genius of great scientists such as Newton is not in determining the mathematical constraints that hold between some existing F and p , but in the description of concrete real world problems in terms of newly invented abstract concepts such as force and momentum – simplifying calculation in the process. Expressions in terms of p and F are meaningful when the reader understands how to recognize what things in the domain of interest these concepts apply to, and how to measure mass m , time t , and velocity v .

The same thing applies in the legal field: legal concepts must eventually be defined in terms of concepts that can be operationalized in the domain of interest, by the people who should understand and apply the law.

A *knowledge representation language* is a language designed for the purpose of knowledge representation. Not all data structures are equal from a reasoning perspective: Representing the knowledge in one way may make the solution simple, while an unfortunate choice of representation may make the solution difficult or obscure.

Long division is for instance much simpler with Arabic numerals than with Latin numerals.

Copernicus' claim to fame is in showing that assuming that the sun does not move is a whole lot easier than assuming that the earth doesn't move, if one wants to predict the movements of celestial bodies. Interestingly, both assumptions are clearly untrue: both sun and earth move relative to the rest of the universe.

Most practical physics problems are easier to solve with Newton's model of physical reality than with Einstein's, even though Einstein's is considered a better fit to reality in marginal cases. Newton's theory is obsolete as an explanation of how the world works, but it still *works* as an instrument.

Different knowledge representation languages are designed for different domains of interest. The field of legal knowledge engineering designs its own languages. The reason for having so many different representation languages is that these languages are tailored for expressing specific relevant aspects in the domain. This special focus gives them their expressive power and helps the knowledge engineer in focusing attention to the right things.

This book limits its attention to knowledge representation used by KBS. The whole of knowledge representations used by some KBS is its *knowledge base*.

Most knowledge representation languages are *logics*. Logical systems are formalizations of the proper methods of reasoning and logical truths are those demonstrable by these correct methods. In this sense logics are a special kind of prescriptive theories of proper reasoning. The key notion in logic is the notion of a valid argument. An argument is valid if and only if, if its premises are true, then its conclusion is true. An argument is sound if and only if it is valid and it has all true premises. The logic cannot tell us whether the premises are true: we are the judges of that.

An automated reasoning method for deciding arguments is *adequate* if it decides that an argument is valid if and only if the argument is valid. It is *complete* if it is adequate for all arguments that the language can express. It is *decidable* if it is possible to decide on the validity of any argument in a finite number of well-defined steps, and *tractable* if a computer can produce such a decision in a reasonable time frame using the reasoning method.

The *expressiveness* of language is the extent to which the language allows us to represent the knowledge we want to express. Knowledge representation languages usually differ in expressiveness, and are geared to some type of problem. Knowledge representation languages do not always differ in expressiveness, even if they are designed for different purposes. Logicians are often of the opinion that two knowledge representation languages that are equally expressive from the logical point of view, are really just the same language written in another way. For knowledge engineers this is not the case: languages designed to be used for representing different things cannot be considered equally, more, or less expressive relative to each other: ordering them on the same expressiveness dimension is a non sequitur.

The claim of logical reasoning is that, whatever truth may be, it provides methods to make sure that conclusions are as true as the premises, if the axioms are true. A logic is *monotonic* if new facts cannot invalidate old conclusions, i.e. change the truth value of a proposition. Monotonic reasoning does not allow for exceptions. If something is believed to be true, we will continue to believe it is true regardless of any new information we gather about the world.

Monotonicity and tractability are very appealing properties for a knowledge representation language to have. The great advantage of monotonic reasoning is that the order in which inferences are carried out does not affect the conclusion of reasoning. In other words, we do not have to worry about ordering effects in the knowledge representation language, the order in which knowledge representations are added together in a knowledge base, or the order in which facts and claims are discovered or supplied. The property of monotonicity, in combination with tractability, has special significance in the context of reusability (see section 2.2.4) and the Semantic Web (see section 3.2). Monotonicity is a good *design principle*, because it helps to decompose systems into independent components, even if it is a bad model of how human beings solve problems.

In the legal field it has become common to use nonmonotonic logics, because *isomorphism* between the defeasible reasoning policies expressed in the sources of law and the knowledge representation appears to demand it, and because of the essentially adversarial nature of legal argumentation. In section 2.5 I will explain part of the case for defeasibility in law.

Isomorphism can be characterized in different ways, but the most obvious one is that a human being should be able to understand the relation between the knowledge representation and a source of knowledge being represented. Isomorphism does not just apply to for instance the mapping between legal knowledge sources – generally writings – and knowledge items, but also for instance to justification: the human observer should be able to understand how the system reached the conclusion it reached.

Isomorphism as used in legal knowledge engineering is suggestive of the more precise term used in mathematics and the natural sciences. Isomorphic structures are *structurally* identical at a certain chosen level of granularity. This level of granularity may in the case of written sources be the word, the sentence fragment, the sentence, the complete or article, etc.

Isomorphism between the source of law and its knowledge representation (introduced in this sense in [17]) is an objective of this book. Isomorphism with argumentation and justification is a more contentious issue, as section 2.5 explains.

Although most knowledge representation languages are logical languages, their properties cannot be fully reduced to the properties of the associated logic. The language comes with *ontological commitments*, as explained in the next section.

This book does not propose or recommend a specific logical language. As stated in chapter 1, use of the OWL DL language is a design requirement for this project, but this book is about representation choices and their practical consequences for the management of complex, version-managed, and distributed knowledge bases.

2.2.1. Knowledge Representation and Meaning

Reasoning is a process that goes on inside the mind, while most of the things we wish to reason about exist only in the outside world. The knowledge representation is a stand-in for the things that exist in the external world. Operations on and with representations substitute for operations on the real thing, i.e. substitute for direct interaction with the world. Reasoning itself is in part a surrogate for action in the world.

Viewing representations as *surrogates* leads naturally to two important questions.

The first question about any surrogate is its intended identity: what is it a surrogate for? The correspondence between surrogate and the real world referent is a fundamental component of the meaning of the knowledge representation, also sometimes called the *extension* or *extensional semantics*, not to be confused with the operational or denotational semantics of the underlying logic.

The second question is how close is the surrogate to the real thing? This is the *fidelity* of the knowledge representation. Perfect fidelity is in general impossible, because any thing other than the thing itself is necessarily different from the thing itself. Fidelity can however always be increased by investing more effort in the representation.

Note that the identity and fidelity question apply to the things we are reasoning about, while the isomorphism question addressed before applies to the information we choose to use as a source of knowledge.

Knowledge representation serves equally well for intangible, *mental things* such as thoughts, moral debts, and plans as it does for tangible, *physical things* – whose

existence can be experienced through the senses – such as gear wheels, centers of gravity, or clouds, allowing humans and KBS to reason about both types of things, and their interaction. Somewhat paradoxically it is even perfectly normal to use something to represent nothing³.

Abstract things, which should not be confused with mental ones, can of course exist inside the KBS with *perfect fidelity*: Mathematical sets and graphs, for example, can be captured exactly, precisely because they are abstract objects. Abstract objects are objects that exist only because someone declared them: it is in principle possible to state the necessary and sufficient conditions of abstract terms, although one may choose not to do so. There is in fact no appreciable difference between the thing, for instance the set {1, 2, 3}, and its representation {1, 2, 3}, and no difference between the extension of set theory and its denotational semantics. Something is abstract by virtue of having no extensional meaning different besides its denotational semantics.

One could argue that denotational semantics is a kind of intermediate form between the messy outside world and the mind. Since almost any real world reasoning task will encounter the need to reason about things outside as well as abstract things, imperfect surrogates are however inevitable.

The distinction between mental and abstract things is described in more detail by us in for instance [153, 154]. The notion of the mental thing is borrowed from Dennet's *intentional stance* in [95]: the world of mental things is an alternative route to explaining physical events, applied to the behaviour of human beings and, to a lesser extent, animals as actors. The mental category subsumes epistemic and intentional things: epistemic things are mental things that have to do with (acquiring) knowledge. Intentional things only exist because we intend them. It is of course possible that one represents one's own mental state, and (at least) a KBS can do so with perfect fidelity. Whether humans are good at uncovering their own mental reality through *introspection* is a point of contention (cf. for instance [191, 257]).

The knowledge representation itself, if one for the moment forgets that it is *about* something, behaves like an abstract object, and the properties of abstract objects can be used to predict the unknown properties of the referent. It is for instance possible to use a theoretical disk, with an evenly distributed weight, to predict the center of gravity of a porcelain plate. The real center of gravity of the porcelain plate is however observable: it is the place on which it can be balanced. The belief that the center of gravity is equal to the center of gravity of the idealized object is a *hypothesis*, that can be falsified by observation. This hypothesis is undoubtedly useful, because porcelain plates are very breakable: it is a good idea to first predict where it is before one tries balancing it. This predictive quality of the knowledge representation, through the denotational or operational semantics of the representation, can be called it *intension*, or *intensional semantics*.

One of the attractive features of the law is its similarity to knowledge representation. There is no issue of fidelity with respect to outside referents as long as the processes of applying the law are not taken into account. The fidelity question is reduced to the

³In OWL the distinction between `owl:thing` and `owl:nothing` is in fact the most fundamental ontological distinction made.

fidelity of the copy of the sources of law into the knowledge representation language, which is unfortunately generally less expressive than the language the legislator uses. In reality things are of course slightly more complicated.

Law deals with the same fidelity issue as knowledge representation: the correspondence between legally declared reality and social reality can break down. Generally speaking the legislator explicitly decides to leave the interpretation of certain terms used in legislation open to interpretation. In this case legal theory often speaks of *open textured concepts* (cf. [237]). In its broadest sense all physical and mental things described in legislation, its regulated domain, are open textured. In a more limited sense, a concept is called open-textured when the legislator is being *intentionally* vague. Open-texture is for instance used if the legislator believes that it is not practical to try to narrowly define terms because the regulated domain of interest⁴ changes too often in unpredictable directions or is hard to pin down in general (cf. [284, 237]). In other cases, the use of open-textured concepts is a compromise, the consequence of disagreement in group decision making.

Two important consequences follow from the inevitability of imperfect surrogates. One consequence is that in describing the world, we must inevitably resort to untruth, by omitting details at least. At a minimum we must omit some of the effectively limitless complexity of the outside world. In addition, our descriptions may introduce artifacts not present in the world.

The most important consequence is that reasoning about the world must eventually reach conclusions that are incorrect, independent of the quality of the reasoning process used, and independent of the quality of the representation used. Logically sound reasoning cannot save us: If the world model is somehow wrong (and it must be) some conclusions will be incorrect, no matter how carefully drawn. A better representation cannot save us: all representations are imperfect and any imperfection can be a source of error.

If all representations are imperfect approximations to reality, each approximation attending to some things and ignoring others, then in selecting any representation we are in the very same act making a set of decisions about how and what to see in the world. Selecting a representation means making a set of *ontological commitments*.

These commitments and their selectiveness are not some unfortunate and incidental side effect of a representation choice. A knowledge representation *is* the realization of a set of ontological commitments. It is unavoidably so because of perfect fidelity is impossible. It is usefully so because judicious selection of commitments provides the opportunity to focus attention on relevant aspects of the world, while ignoring the infinite details. Just like the court, the knowledge representation only admits relevant evidence.

Part of what makes a language a representation language is that it carries extensional *meaning* or extensional semantics, i.e. there is a correspondence between its constructs and things in the external world. That correspondence in turn carries with it *constraints*. While every representation must be implemented in the KBS by some

⁴Note that I use *domain* in two senses throughout this book: the domain being represented by the knowledge representation and the domain being regulated, and therefore represented, by legislation. The domain represented by the knowledge representation includes the relevant legislation.

data structure, the representational property is in the correspondence to something in the world and in the constraint that correspondence imposes.

If you say in the ontology specification of your domain that all cars are vehicles, this is an ontological commitment, meaning that the system may derive that something is a vehicle because it is a car but for instance not that something is a car because it is a vehicle. This is a constraint on the inferences warranted by the knowledge representation. A system that ignores these constraints on the knowledge representation ignores the semantics of the knowledge representation.

In Davis, Shrobe, and Szolovitz (viz. [88]) knowledge representation was characterized as follows:

1. A knowledge representation is a surrogate for the thing itself, used to enable an entity to determine consequences by thinking rather than acting.
2. It is a set of ontological commitments, i.e. an answer to the question: In what terms should I think about the world?
3. It is a fragmentary theory of intelligent reasoning, expressed in terms of three components: (i) the representation's fundamental conception of intelligent reasoning; (ii) the set of inferences the representation sanctions; and (iii) the set of inferences it recommends.
4. It is a medium for pragmatically efficient computation; It organizes information so as to facilitate making the recommended inferences.
5. It is a medium of human expression, i.e. a language in which we say things about the world.

These characterizations position knowledge representation against several dimensions on which the representation can score better or worse; Often trade-offs are involved. Davis, Shrobe, and Szolovitz use the term inference in a generic sense, to mean any way to generate new expressions – or beliefs when talking about a human being – from old ones. Problem solving can somewhat trivially be characterized as making inferences recommended and warranted by the knowledge representation until the problem is solved. The quality of the knowledge representation is amongst others determined by whether a solution is reached at all, and how many inferences are needed to do it.

Additional standards for determining the quality of a knowledge representation can be for instance found in (cf. [120]).

When considered as a reusable resource for production, a good knowledge representation:

1. will allow the reuse of the same knowledge for different types of problems and with different reasoning methods,
2. is easy to maintain when the outside world changes, and
3. facilitates isomorphism between representation and consulted knowledge sources.

Finally, it should be understood that a KBS, and the knowledge base it depends on, embodies a certain *problem solving competence*, i.e. the competence to solve a certain type of problems (cf. generally [187]). A KBS that was built to, for instance, determine income taxes, need not be competent to determine income. The mere fact that it knows about *income* – i.e. there are descriptions of the domain in its knowledge base involving *income* – does not mean that it *knows enough* about income to help

determine it. This distinction is central to the design of KBS: the KBS may embody an ontological commitment to a certain conception of income, but at the same time may lack operational knowledge about determining it.

2.2.2. Limitations of Symbolic Knowledge Representation

A popular objection to knowledge representation – at least of the symbolic or “logical” approach – holds that knowledge representation is not a viable model of human intelligent behaviour at all. Proponents of this thesis deny that the brain can be conceived of as a linear processing machine using “knowledge” stored in a memory, and advocates alternative processing models, for instance neural networks, that offer a better analogy to what we know of how brains really work. This discussion of course relates more to philosophy of artificial intelligence, something the field of legal knowledge engineering has never really been concerned with⁵, than to the utility of knowledge engineering in the design of decision support systems and knowledge-based systems.

The neural network approach has not gained currency in the legal field. One of the problematic aspects of this approach is for instance articulated by Kowalczyk (cf. [177]): the neural network approach is fundamentally deficient for many uses because of its inability to justify the results reached. If there is any similarity between neural networks and brains, it is definitely lost on *minds*. The law is not just about the decision, but also about the correct justification of the decision, based on explicit knowledge of the legal sources to be applied, and the features that are and are not to be taken into account. Neural networks are not very good at justifying themselves, and cannot distinguish between spurious and relevant correlations between input and output. They are not transparent, and engineers like transparent systems, not just in law. This limits their usefulness to solving problems that humans do not understand well (such as detecting sea mines in sonar data, “electronic noses”, and data mining).

Some have since demonstrated generation of justifications from neural networks (e.g. cf. [128, 16]), and made interesting observations about the conditions for successful application of neural networks and statistical methods to law (cf. [161, 16]). Kowalczyk’s point is however not really about the abilities of neural networks, but about a design philosophy of “training” a general-purpose architecture with a set of cases from the past to achieve competence in producing the same output as the legal system does, even though explicit jurisprudential theories aiming at the same thing exist. The inputs and outputs of the system have to be so carefully engineered to be justifiable for application in law that it matter-of-factly becomes symbolic knowledge representation, or as John Sowa put it, on his first encounter with a call for papers on *feature engineering for machine learning*⁶:

For years, the people who were working with neural nets and other learning systems were hyping the claim that they did away with “knowledge engineers” because they didn’t need rules, frames, formulas, etc. But now they discovered that these learning systems don’t really learn all by themselves – somebody has to analyze the subject domain and select

⁵A variation on a well known epigram: The philosophy of artificial intelligence is as relevant to artificial intelligence as ornithology is to birds.

⁶Email archive: <http://suo.ieee.org/email/msg12963.html>

some set of “features” for the learning systems to work with. Since they had already got rid of the knowledge engineers, they now discovered a totally new breed of “feature engineers”.

Law and knowledge engineering both operate with “folk psychology” cognitive models of human knowledge, intentions, and behaviour that are incorrect according to postmodern critics armed with scientific knowledge from psychology and neurology (viz. [6]). This resort to “folk psychology” is for instance very evident in the ascription of intent to suspects of crimes, based on their observed behaviour, and assumptions about what they must have foreseen as the consequences of their actions, given what they must know. In chapter 4 I will touch on this point again and discuss its relevance – or rather its lack of relevance – for legal knowledge engineering. See for instance [277] for a discussion, in the context of a critical evaluation of the role of expert witnesses in court, of the differences between effective and justifiable decision making in law and empirical science as a quest for the truth.

The law applies a set of assumptions about human behaviour very similar to what knowledge engineer Allen Newell (cf. [209]) called the *knowledge level hypothesis*. The knowledge level permits predicting and understanding behaviour without having an accurate operational model of the processing that is actually being done by the agent. Observers treat an intelligent agent as a system at the knowledge level: they ascribe knowledge and goals, and a variety of other *mental things*, to it. According to Newell knowledge can be defined as “whatever can be ascribed to an agent such that its behaviour can be computed according to the principle of rationality,” noting that the principle of rationality means that if an agent has knowledge that one of its possible actions will lead to one of its goals, then the agent will take that action.

Newell’s hypothesis does not exclude the possibility that “knowledge” is best captured in some kind of neural networks, but by stressing that knowledge is something that is “ascribed” by the outside observer (i.e. it is an *externalist* and *doxastic* epistemological theory of behaviour) it clearly leads to the conclusion that descriptive fidelity in a neurological sense is much less important than more standard desirable properties of a scientific theory such as:

1. its effectiveness as a predictor;
2. its value as a justification; and the
3. ease of communication of the model.

The third and second functions are clearly better served by a symbolic knowledge representation. Also, by assuming the “principle of rationality” as a constant in the prediction of behaviour, it must necessarily follow that the fact that agents do different things in the same circumstances must be explained in terms of different knowledge bases. The claims the “principle of rationality” makes about human behaviour are in effect quite modest: it makes no assumptions about the processing equipment, and even declares itself invulnerable to falsification by introspective accounts.

In legal knowledge engineering, more specific charges have been made against the logical approach. It has often been claimed that legal reasoning – mostly by theorists from outside the field of legal knowledge engineering – cannot be captured by a logic. These criticisms boil down to three types.

The first one is a mere restatement of what we already noted before: that the real world cannot be captured with perfect fidelity. Any knowledge representation, including legal knowledge representation, will at some point sanction inferences that do not lead to true conclusions. When you are making a new design plan for your garden, the theory that the earth is flat is perfectly viable. But when you are traveling to another continent by ship, the theory breaks down, and has to be replaced by the theory that the earth is round. This is also true for an LKBS (see for instance [201]): even if it works in most cases, at some point the LKBS will make a decision that perhaps a human would not have made because the human is grounded in reality in a way that the LKBS is not.

This criticism is acknowledged in knowledge engineering, and is the main reason for the interest in making explicit the ontological commitments a knowledge-based system is based on. The decision of the LKBS is as much a prediction as for instance a lawyer's assessment of what a court would decide. Another consequence is obviously that the responsibility for maintenance of the LKBS and the liability for the decisions it takes should be critically evaluated when a LKBS is introduced.

The issue of *falsifiability* of the conclusions of the LKBS may also be connected to defeasible reasoning. One should however keep in mind that also an LKBS based on a knowledge representation language that allows for defeasible reasoning will at some point sanction inferences that do not lead to true conclusions. The issue here is one of isomorphism rather than of fidelity of representation.

The second criticism is more specific to legal knowledge representation: legal knowledge representation languages usually prominently represent norms. The usual presentations of deontic logic, the logic of norms, whether axiomatic or denotational, treat *norms* as logical sentences that bear truth-values. A fundamental problem of deontic logics, is to reconstruct it in accordance with the philosophical position that norms prescribe rather than describe, and are neither true nor false. In other words, norms do not have truth values. This observation is generally attributed to mathematician and philosopher Walter Dubislav (viz. [103]).

This criticism may be waved away by noting that concepts, triangles, planets, and sets do not have truth values either. Only the statements *about* them do, by affirming or violating the ontological commitments we made regarding these objects.

One may for instance say that such and such a norm is (or is not) part of (or implied by, consistent with, etc.) such and such a normative code. For example one may say that the British driving code requires vehicles, in normal circumstances, to be driven on the left hand side of roads, while continental European codes requires them to drive on the right side. These statements report on the presence or absence of a norm with a certain effect in a given normative system, and are often taken to be the fundamental representation category in deontic logics. To mark the difference, they are sometimes called *normative propositions* or *normative statements*.

As suggested for example by Bengt Hansson in [136], it is enough to take one's preferred possible worlds semantics for deontic logic and reinterpret its deontic components as relativized to a given normative system to solve the problem. One may also state that such and such situation or behaviour is a violation of a norm, or that this norm is in conflict with that norm, or that some norm directs us to choose this

alternative over that alternative in a certain situation. In each case we make a statement about the norm, instead of representing the norm. What we however cannot do is equate some proposition or logical sentence to a norm.

As Makinson notes in [195], most people in AI & Law do accept the theoretical problem, but in practice work goes on as if the distinction had never been heard of:

In axiomatic presentations of systems of deontic propositional logic, the truth-functional connectives “and”, “or” and most spectacularly “not” are routinely applied to items construed as norms, forming compound norms out of elementary ones. But as [103, 163] already observed, if norms lack truth-values then it is not clear what could be meant by such compounding. For example, the negation of a declarative statement is understood to be true if and only if the item negated is false, and false if and only if the latter is true; but we cannot meaningfully formulate such a rule for norms.

He notes an unfortunate consequence of lack of attention for the normative system as a whole, which leads us back to the earlier point about fidelity: the major problem that plagues legal knowledge engineering is that legal knowledge representation fails to *scale* because we cannot reliably break it down into reliable components. Legal knowledge representation usually works for single knowledge-based systems tailored towards solving a specific type of problem, but breaks down if we put the knowledge bases of two knowledge-based systems together in one logical reasoning system. One of the reasons is the lack of attention for how and when normative statements are composed into larger composite normative statements, if it cannot be reliably done with truth-functional connectives, more specifically an implicit conjunction of normative statements.

In this book two design approaches are used to address this problem. The first one is to consider normative statements as statements *about* a norm; The norm is an object in the knowledge domain, and *not* a logical sentence. The second one is a skeptical attitude to composition of norms into normative systems: the violation of a norm in principle stands by itself, unless our knowledge tells us otherwise.

A third criticism is for instance voiced by Moles in [201]: Legal knowledge engineering focuses too intensively on representing the sources of law instead of the meaning of law. This criticism is somewhat addressed in section 3.5.1. The reader however has to keep in mind that this work is indeed only about sources of law.

2.2.3. Problem Solving and Problem Solving Competence

The engineers of a knowledge based system try to provide a specific intended functionality in an efficient manner. They take into account resource limitations on for instance time and memory of the computer, the availability of information and the cost of collecting information, the required performance of the KBS, and make *assumptions* about the institutional context of the problem solving task, about the availability, quality, and completeness of certain kinds of information, about the knowledge and needs of the users of the system, etc.

This focus on functionality and efficiency is not only necessary because knowledge representation is a resource consuming task, but also because otherwise the design process would be underspecified, missing an explicit goal, and there would be no way to determine that the result is acceptable.

Knowledge representation is usually only suitable for a specific *problem type* that the knowledge engineers had in mind. Each problem type supplies its own view of what is important to attend to, and each suggests, conversely, that anything not easily seen in those terms may be ignored.

A KBS designed for buying and selling second-hand cars will generally speaking not be very helpful for someone designing a knowledge-based system for diagnosing and repairing cars or driving cars, even though all three contain descriptions of cars: the required information about cars is very different (cf. [153] for this example).

In terms of the porcelain plate example used earlier: A KBS that helps us to determine the center of gravity of porcelain plates usually does not explicitly communicate its theory on idealized porcelain plates (i.e. as disks with an evenly distributed weight), and has no knowledge at all that would help it plan the production of a porcelain plate.

A *problem solving method* (cf. [62]) supplies a coherent set of assumptions about what to attend to and what to ignore to achieve a certain *epistemic competence* or *problem solving competence*, i.e. to be able to solve a type of problems. Fensel and Straatman in [109] for instance characterize the problem solving method by a functional specification, an operational specification, a cost description, and a set of assumptions about the domain and the available domain knowledge that must be true for the problem-solving method to realize its functionality.

Note that more often than not we make assumptions that are known in advance to be not true, but true often enough to be practical, in order to reach conclusions. This is for instance often the case with the *closed world assumption* (cf. [251, 205]), which states that we know all that is relevant to know, or *frame assumptions* (cf. [86]), which state that no unforeseen outside processes affect the state of some relevant object or situation, and which are often used in planning problems. An assumption specific to legal planning is for instance that other agents do not generally violate the law. This assumption makes for instance driving in busy traffic workable: we do not have to consider all possible ways in which other traffic participants could decide to violate traffic law.

The use of unverifiable assumptions, certainly if they are left implicit, means that a knowledge representation suited for one type of problem – for instance planning – is not automatically suitable for another type of problem – for instance diagnosis, not even if we add relevant knowledge to it.

2.2.4. Reusability and Knowledge Components

There are good reasons to expect that knowledge representation that is neutral with respect to problem solving competence is in principle impossible (cf. generally [76]). For one thing, such a design problem would be underspecified as pointed out before. Secondly, one could argue that knowledge, as understood in the Newellian sense ([209]), simply cannot be disconnected from the problem solving competence it helps to realize: we call some data structures knowledge representation by virtue of the problem solving competence they help to realize when properly understood by a rational agent. This is a necessary consequence of the external position we take towards intelligent agents when we ascribe knowledge to them: the agent first has to show it is competent to

solve a certain type of problem before it can be said to be knowledgeable.

The degree of reusability of the knowledge base as a whole has already been fixed by the designers of the system: the knowledge base is reusable only relative to that part of the knowledge base that varies per problem solving session. We can call this variant part *the case*, and the rest the *invariant knowledge base*, or – if no ambiguity exists – simply the knowledge base.

In the case of LKBS, the case usually does not *include* the sources of law, even though these are subject to change. Since the sources of law change, in some domains regularly, and moreover usually change outside of the control of the users of the LKBS, the LKBS maintenance task is both labour intensive and hard to schedule resources for.

Regular maintenance imposes some constraints on the design of LKBS. Firstly, it must be easy to identify *where* the LKBS must be changed in response to a change in the sources of law. This is the problem of *isomorphism* between source of law and representation. Secondly, we want to change them in only one place. Even better is if we can make a single change to *all* deployed LKBS that use the same source of law.

We therefore compose the knowledge base in as far as possible from reusable *knowledge components*, that are maintained at a single point. This is the notion of *knowledge syndication*. The component in effect becomes a generic knowledge base interrogated by multiple specific purpose LKBS, which creates a coordination problem: we must make the change in such a way that all user KBS keep using it correctly.

Focus has changed over time from knowledge as the attainment of truth to knowledge as productive factor (cf. [191]), to be realized as – one-shot or reusable – plans or – reusable – designs for machines⁷. The first perspective is often called the (traditional) *philosophical* perspective, while the second can be characterized as the *economic*, or more generally the *social sciences* perspective. It is this shift that introduced the concept of *tacit knowledge*, the knowledge that bridges the epistemological gap between explicit “theoretical knowledge” that explains certain well-understood aspects of a situation and *skill* on the work floor (cf. for instance [191]), which involves the less well-understood capacity of intelligent *situated action*. The common (mis)conception of tacit knowledge as knowledge that is hard to make explicit – alluded to in section 2.1.1 – is only secondary to this first meaning: if tacit knowledge is a stop-gap concept, referring to those aspects that are not well-understood, it ceases to be applicable as soon as the situation is well-understood.

The study of reusable knowledge components is in essence about knowledge as a factor in the production of yet more knowledge, the next shift in focus as evidenced by the *Semantic Web* vision, discussed in section 3.2. Interestingly, this new shift has led to renewed interest in the philosophical perspective.

While it may be impossible to make a representation problem independent, we can restate the issue to enumerating and describing types of problems and problem

⁷[191] uses classical economic terminology: capital, labour, plans, and machines. Plans and machines in this context evoke business processes and conveyor belt-based assembly-lines. Tool, device, or artifact are other possible terms for what I mean here: I interpret this as anything manmade that requires non-trivial knowledge to make and has some *function* or purpose, ranging from computers to hammers and violins.

solving methods, and understanding the systematic relations between them and their components, in the hope that this gives us clear ideas about what constitutes a reusable knowledge component. Attempts to make such an inventory, for knowledge engineering in general, are found in for instance [59, 206]. The other, “philosophical”, avenue of exploration, exemplified by top ontologies (viz. [186, 185, 255, 222, 197, 114], simply divorces the search for fundamental truth from problem solving, in the expectation that the truth part is reusable, and is of less direct interest for this discussion.

Breuker in [59] (see also [33, 34]), distinguishes the problem types modeling, planning, design, assignment/scheduling, monitoring, assessment, diagnosis, repair, and remedy, and postulates a number of functional dependencies between them, observing that the output of one problem type is often the input of another one. The problem types impose different epistemological perspectives on the same domain (knowledge).

Design is for instance characterized as the composition of building blocks with a certain functional interface specification into a system specification meeting certain constraints and requirements. Diagnosis takes observations of the behaviour of such a system, that is observed to deviate from its specification, and determines which building block(s) does not behave according to its specification.

Breuker’s list of problem types clearly bears the marks of the focus on knowledge as a productive factor, a resource: design, diagnosis, and repair revolve around machines that are supposed to have some function or purpose, while planning, scheduling, and assessment revolve around a plan that attains some goal.

Complementary to the problem types approach are attempts to automatically design problem solving methods from components, as proposed in for instance [264].

The problem types are surprisingly versatile in their mapping to domain knowledge: design applies to such varied things as violins, organizations, computer programs, mathematical proofs, business processes, etc. The mapping is of course not completely free of constraints: you cannot plan a train, but you can plan to take it, or plan its construction.

In the context of legal knowledge engineering, planning (subject to the constraints of the law) and assessment (of cases) appear to be the legal problem types par excellence (as noted in for instance [265, 275]). This distinction is in itself fruitful, as it for instance helps us to realize that the *ex ante* role of the norm in planning is very different from its *ex post* role in recognition and assessment after the fact (cf. section 4.7). This type of differentiation of role in problem solving is called the *epistemic role*, or simply the knowledge role, of knowledge items.

It is however also easy to come up with typical design or diagnosis scenarios in the law. In my own work reported in for instance [284], the domain of interest is firstly ship design, subject to legal constraints and requirements, and only secondarily ship handling. One could also consider typical form filling systems (electronic tax filing, legal drafting) to be typical parametric design systems.

Deciding which driver(s) caused a chain collision can for instance be conceived of as a form of diagnosis, if we accept the machine metaphor for traffic flow, where we apply typical diagnostic assumptions such as the *single fault assumption*: i.e. we first consider all hypotheses assuming some non-standard behaviour of one component, while all others are assumed to behave according to specification, and see whether one

of these hypotheses explains the evidence. Only when this doesn't yield a conclusion we proceed to the assumption that two components failed, etc. The drivers are the metaphorical components.

The conclusion must be that the field of law is *not* strictly characterized by attention to specific types of problems: how it conceives of legal problems is determined by the nature of the problem and the character of our domain knowledge. We cannot establish a systematic relationship between the *domain* of law and specific problem types, and there is no apparent reason to expect that there is such a relationship.

Separation of reusable knowledge about the domain and reusable knowledge about problem solving, or ontology and epistemology, as characterized in [26, 153, 154], and in this book, is however the first step to bridging the *epistemological gap* and designing reusable knowledge components. This proposed solution comes with a cost: reusing existing components, and designing knowledge components for reuse, is certainly initially more costly than designing a solution for a single problem, and this increased cost must be justified in terms of reusability, maintenance, or scalability.

This distinction between knowledge of the domain and knowledge of problem solving has gained some currency in knowledge engineering in general, but the distinction between the ontological and epistemological aspects of knowledge as a productive factor is not often explicitly made in legal knowledge engineering⁸, where the need for separating knowledge by epistemic role is seldom felt. An analysis of legal reasoning in the context of this distinction is therefore in itself worthwhile. Moreover, if law is not characterized by its own problem types, it should be possible to deal with legal reasoning using standard knowledge engineering techniques.

Another important methodological admonition that can be gleaned from this section and the preceding one is that a KBS should explicitly characterize the problem solving competence it is intended to realize in its knowledge base, so that we can at least assess the value of the knowledge it represents as a productive factor beyond the immediate context of use for which it was designed.

Moreover we should distinguish between the porcelain plate and the theoretical disk of section 2.2.3: this distinction involves the concept of abstraction.

2.2.5. Ontology and Epistemology

When considering knowledge representation, we can distinguish issues of ontology – about the *object* of our knowledge – and issues of *epistemology* – about the processes by which knowledge is attained. This is not the traditional philosophical distinction. In fact, the sense in which we use epistemology is closer to what cognitive scientist Christopher Longuet-Higgins⁹ named *epistemics* – to distinguish it from epistemology – instead of epistemology proper. The distinction is close to the language level distinction between the extension of the sentence, and its semantics proper, but even here we have to distinguish between the semantics brought in by the ontological commitment and the semantics brought in by what we try to do.

⁸[202] is an exception, as well as the work of [153, 154] mentioned earlier.

⁹The first chairman of a “school of epistemics”.

In this book I will often use the adjective *epistemic* when discussing particulars (e.g. a hypothesis is an epistemic object), and *epistemological* when discussing universals (e.g. hypotheses have a lower status relative to truth than facts). For ontology I can make no such distinction, as *ontic* would be considered a confusing neologism by the reader.

So far I have mostly discussed the semantics of knowledge representation in relation to ontological commitment and fit with an outside world. Epistemology imposes another criterion: to know something is to have acquired the belief in that something through some process by which knowledge is reliably achieved. Note that epistemic reliability cannot be simply reduced to a statistical criterion (see for instance [202]: p. 42–43); A process that more often than not produces a true answer, does not necessarily discriminate correctly, need not be operating properly, etc.

Epistemology and ontology of course cannot strictly speaking be separated: production of knowledge is obviously inextricably bound to ontological commitments.

The distinction is another one: the separation of the conceptualization of the “outside” world the reasoner is acquiring knowledge about, the *knowledge domain*, from the conceptualization of that part of the mental world of the reasoner that plays a role in acquiring knowledge. From an ontological viewpoint there is nothing that relates bin packing and scheduling lectures to classes and time slots: from an epistemological viewpoint both are assignment problems. Epistemological and ontological knowledge are linked by a non-obvious mapping, usually handmade by the knowledge engineer, that, if certain assumptions are met, gives a reasoner that understands it a certain epistemic competence, i.e. problem solving competence, in the domain covered by the ontology.

The distinction between epistemology and ontology in this sense helps reuse by separating two aspects of knowledge that can be reused under different conditions. Epistemological knowledge can be reused in different knowledge domains if we are solving essentially the same type of problem using the same types of knowledge sources, while the same ontology can be reused for problems directly or indirectly involving the same knowledge domain.

For the subject of this book this means that we are interested in ontological knowledge about the world of law, as expressed in the sources of law. Epistemological knowledge plays a role to the extent that the sources of law themselves model that mental world.

2.3. Ontology

In philosophy *ontology* is the study of *the nature of being, reality, and substance*. The *ontology* as it is usually understood in Knowledge Engineering has a more specific meaning: it is the product resulting from the systematic inventory by knowledge engineers of relevant aspects of a certain knowledge domain. This necessary restriction of attention to *relevant* aspects relates to the ontological commitments of Davis et al. (cf. [153, 154, 88, 67]).

Ontology in the knowledge engineering sense is often defined as “a specification of one’s conceptualization of a knowledge domain.” (cf. [131]). The specification is

usually made in a knowledge representation language, but it is important to distinguish between the specification (which is usually called the ontology in practice) and the conceptualization, which embodies the ontological commitments and of which it is a specification.

Secondly, ontology is also often described as a *shared* conceptualization of a domain, which stresses the prescriptive nature of the ontology as a kind of quasi-contract binding the developers of the KBS and its prospective users, and a vocabulary they can use to communicate with each other. The same concept can also be applied to interoperability between systems. The ontology often also functions as a schema or specification against which messages can be validated. Open standards in IT are for instance often said to have a ontology (cf. generally [53, 182, 245] for this purpose).

The ontology specification usually consists of terms T , denoting a set of things to which the term can be applied, to be specified by terminological axioms of the form $T \subseteq C$ (C is a necessary condition for T) or $T \supseteq C$ (C is a sufficient condition for T) specifying them. In rare cases a term can be *defined* – all necessary and sufficient conditions can be supplied – by an axiom $T \equiv C$; This is most common for abstract terms, for reasons noted earlier. The terms *term* and *concept* are used interchangeably for most practical purposes. I generally prefer term for a named concept: T is a term, while $T_1 \cap T_2$, occurring at one of the sides of a terminological axiom, is a concept (the intersection of the extensions of T_1 and T_2) but not a term. The terms together form a terminology or vocabulary, and in combination with the terminological axioms the ontology.

It is important to note that each and every statement in a knowledge representation carries with it ontological commitments, but that only those statements that are considered *terminological* in character are part of an ontology specification proper.

For instance: that a father is the male parent of a child is ontology, and contingencies such as that John is the father of Jane, that a father of a child is usually married to the mother of that child, or that a parent is entitled to some child benefit from the government, are not.

Another example, from legal theory: that legal responsibility is the liability to be subjected to a sanction of the law is ontology of law, but that one must have done something in some sense to be legally responsible for it is not, at least if we follow Lehmann in [184], or originally Hart and Honoré in [141].

Because the description of the ontological commitments is stored in a knowledge representation language in the form of terminological axioms, and may become part of the knowledge base of a KBS, there is a lot of confusion between knowledge representation, knowledge bases, and ontologies. It is not uncommon to refer to every knowledge representation, or even the physical embodiment of a knowledge base in the form of a file or database, as an ontology.

This has the unfortunate consequence of bringing all kinds of discussions about knowledge representation issues that have no bearing on ontology into the ontology field. There are in the opinion of the present author no such things as fuzzy, Bayesian, or defeasible ontologies, for instance: this is anathema to the notion of shared vocabulary, and we cannot properly characterize the epistemic competence of a KBS that does not commit to its ontology.

Ontology consists of terminological axioms, and axioms are epistemically speaking self-evident: known to be true by understanding its meaning, its intended fit to reality, without proof or possibility of falsification. Inference on terminology alone strictly speaking doesn't produce knowledge: it merely reformulates what is already believed.

To design ontologies, it is good to follow some simple design principles. Firstly, it is bad practice to try to achieve a given epistemic competence in the ontology by searching for a set of sufficient conditions for key terms. Ontologies and terminological reasoning are only intended to help to:

1. rephrase descriptions of things into other terms, to
2. recognize that one description terminologically subsumes another, or
3. that a description is terminologically incoherent.

Axioms should not be added to the ontology specification for other purposes: these propositions have another epistemic role, and contrary to our absolute commitment to terminological axioms we are willing to retract them if they do not fit reality.

It is good practice to insist that any term used by a KBS should be specified in an ontology: the ontology is the main access point into the knowledge base for determining what a certain term means to the KBS. Note that not only individuals, which are entities that have no instances, but also universals, entities that have instances, can in principle be described by terms (cf. [114]). Most ontologies aim to be ontologies of individuals (but unsuccessfully if we follow the argument in [154]).

It is good practice to see to it that all terms in the ontology are related through terminological axioms with other terms. One should expect to find *coherent* terminologies – islands of terms connected to each other by terminological axioms. Single, unconnected terms identified only by waving in the general direction of some external class of referents are meaningless. Coherent terminologies naturally associate with specific (families of non-terminological) theories of the world. Abstract terminologies should always be associated with some theory: this is the very reason for their existence. There is however no reason to expect that terminologies which happen to concur very regularly in descriptions of domains have an *ontological* connection: the reason for the systematic coincidence may be found in epistemic habits.

An important pragmatic aspect in information exchange between human beings is the aspect of quantity: being informative requires that the information one conveys is *relevant* (cf. [227]). One tries to limit oneself to expressing information that adds to the shared knowledge base. Anything that is obvious from discourse *context* will not be made explicit in written information sources, but should be made explicit in the ontology. Consideration of the relation (and the difference) between ontology and *information* yields some additional pieces of useful advice.

When we base an ontology mostly on written sources, and in legal knowledge engineering we are likely to do that, we need to be aware of “hidden” participants in the context of the utterance. One for instance cannot conceive of a parent without a child (i.e. $\text{Parent} \equiv \exists \text{child} \top$), but information contexts where only someone's status as parent (without explicit consideration of the child) is relevant, are possible. Establishing someone's status as a parent introduces the existence of at least one child *of* that parent in the discourse context, and one may later in the discourse refer to

that child.

One could think of metonymy and synecdoche as overzealous reduction of the quantity of information exchanged. Clearly, *metonymy* and *synecdoche* do not belong in an ontology: one does not allow that a Volkswagen is a type of car instead of a car brand, or that someone pays the taxi instead of its driver. This is a very common oversight, and a major source of variety in ontologies about the same subject.

Very importantly, one should be cautious of allowing *metaphor*, even conventional metaphors (cf. [183]), into the ontology, in particular if they pertain to physical entities. Instead one should ask oneself whether the metaphorical use of language indicates the existence of a useful abstraction, as explained in the next section about granularity.

As pointed out by Lakoff et al. in i.a. [183], terminology for certain basic physical phenomena – for instance physical orientation in space (up, down, etc.) – is readily transferred to other things – physical, mental, social, institutional, abstract – in the form of conventional metaphors.

In the case of mental, social, and institutional things we perhaps cannot avoid this use of conventional metaphor (cf. [225]), as it is arguably (an aspect of) the fundamental mechanism by which we construct these concepts in the first place (according to [183]), and therefore part of their fundamental meaning. The transfer of metaphors at least tells us that entities share something important, but – again – this something may simply be an epistemic habit.

In the case of abstract entities, we could argue that time, orientation, containers, fluids, etc. give us the original archetypes for abstract things such as preorders or sets, which can then be applied to other things. The use of metaphorical language can be considered a hint that things share a conventional abstraction. Contrary to [183] I will make no assumptions about the evolution of conventional metaphors and abstractions.

Given the observations about discourse context and about metaphors, one should be ready to accept that there are things generally referred to by the same name that refer to different concepts, so-called homonyms or polysemes. Polysemes are different, but closely related, senses of a concept. Linguists have a rule of thumb for recognizing them: If a word, that does not appear to be an obvious homonym, seems to exhibit *zeugma* when applied in different sentences, it is likely that the contexts these sentences evoke bring out different, but related, senses of the same word. For instance “the state is free to execute its laws and citizens as it sees fit” or “oh, flowers are as common here, Miss Fairfax, as people are in London.”. This rule of thumb is less useful than it may seem, since finding the right sentences is hardly a trivial exercise.

Metaphor transfer is one likely explanation of the common occurrence of polysemes. Polysemy and homonymy is generally solved in ontology specifications by declaring such different concepts in different “namespaces”, e.g. `law:norm` vs. `measure:norm` vs. `social:norm`, or `finance:check` vs. `chess:check`.

In legal knowledge engineering, ontologies are the primary interface between knowledge engineering and legal theory. This work can be characterized as *computational legal theory*, which studies the same subjects as legal theory but with application in knowledge engineering and knowledge management in mind.

In this book I will roughly follow the distinction made in i.a. [153, 154] between

physical, mental, and abstract things. Social things derive from the recognition that you and others have a mental world, the possibility of exchange of information, and the assumption that objects can be shared between mental worlds. Special attention goes to intentional and epistemic things, which are mental things, and institutional things, which are intentional and social in nature. This distinction between a physical, mental, and abstract world is a form of *ontological stratification*, as explained in the next section.

When one speaks of an *ontology of law*, this may mean that it is a specification of specifically legal concepts, or a specification of the relevant aspects of some knowledge domain from the perspective of some type of legal problem solving, or from legal theory. Section 4.2 takes a position on what things are legal or “legally relevant” (to follow a distinction made by Mommers in [202]): legal things are created by the law, while legally relevant things are recognized by it.

Ontologies of law usually contain many, or even mostly, “legally relevant” common sense terms (e.g. [65, 153]): the ontology describes a shared vocabulary, but it is definitely not intended as, or limited to, a jargon dictionary for the legal domain. Ontology of law is more properly characterized as the study of law’s conventional abstractions.

2.3.1. Granularity and Abstraction

The complete ontology of a KBS need not be a particularly coherent description of the world. Copernicus’ elliptical orbits and Ptolemy’s epicycles can harmoniously co-exist, just like the abstraction of a road to a line for navigation, to a surface for driving, and to a volume for repairing potholes.

There is however a difference between the theories of Ptolemy and Copernicus and the three alternative models of the road: the first pair are simply rival explanations, and the second has been judged clearly superior to the first by history, while the three models of the road are all useful perspectives at different levels of granularity depending on what we try to do. Earlier I pointed out that the round earth and the flat earth, and Newton and Einstein’s theories also happily co-exist because they have different strengths and weaknesses.

The notion of levels of *granularity* used in description depending on what we are trying to do appears in different guises in different contexts. In knowledge engineering we find for instance Hobbs’ work in i.a. [150, 149], who introduced the road example. In science we find for instance the excellent discussion of [84] on *reductionism* and how complexity at one level of granularity of description appears to magically collapse into simplicity at another granularity level of description, for no other apparent reason than that it is a better description of our problem. Dennett’s (cf. [95]) intentional, design, and physical stance are another example. Dennett not only discusses the phenomenon of granularity, but also gives a meaningful rule of thumb for ontological stratification into fixed levels of granularity.

Clancey in [81] describes the use of abstraction, aggregation, and refinement as a generic problem-solving strategy.

When we abstract a road to a line, surface, or volume, this is often indicated in

natural language by metaphor. The line abstraction for instance gives rise to *intersections*, and we do see this term applied to road networks. A term such as *box junction* necessitates a refinement to a surface. These spatial abstractions are by no means the only ones for a road: another common abstraction is the hydraulic or cardiovascular one (congestion, artery).

It is important to note that shifts of granularity are not performed on the thing itself, but on abstract models of the thing. The road, and mathematical line, surface, and volume remain different things on the ontological level (see figure 2.1): the abstraction *ideally* speaking simulates the original as in the computer science technique of *bisimulation*. Abstraction is the relation between the thing and its abstract model. In the context of bisimulation techniques we can also distinguish overapproximating and underapproximating abstractions, as in [223], and we can for instance distinguish idealizing abstractions to explain the use of prototypes in knowledge representation. The thing itself, the road, remains firmly fixed to the ordinary realm of the senses, of the conventional referents.

Most deontic logics for instance (consciously) position themselves as overapproximating abstractions, explicitly making predictions not always observed in the original thing. Makinson's argument in section 2.2.2 for instance points this out with respect to the composition of norms. In chapter 6 we will sometimes explore the fine line between underapproximation and overapproximation in reasoning about norms. Some kinds of reasoning about norms are after all more defeasible than others.

The notion of simulation, as used here, has some similarity with the notion of isomorphism from section 2.2 and fidelity from 2.2.1. Each of these can be mistaken for an identity relation and cause confusion.

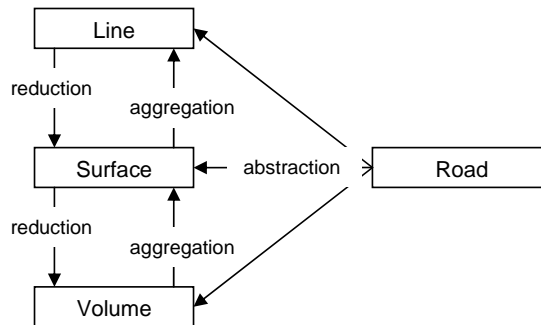


Figure 2.1. Abstraction, aggregation, and refinement.

A refinement moves from a coarser grained abstract model to a finer grained one: it is not the inverse of abstraction, but of aggregation. Abstraction should certainly not be confused with terminological subsumption, the main organizing relation of the ontology, even though it may on occasion be implemented by an operation having the same operational semantics.

Refinement is also not the same as decomposition, although refinement generally involves decomposition (i.e. the refined model has additional parts). Refinements R

are also not by necessity simulations of the aggregated whole A . If it is, we speak of a *perfect refinement* or *reduction* $R \models A$. Refinement is a form of *ampliative reasoning* (to be discussed in section 2.4.1), in the sense that it adds new information not entailed by the problem description. Making predictions about a real thing based on its abstraction is obviously always ampliative.

An unfortunate choice of level of granularity can simply fail to gain traction over the problem. Whether to take a particular stance, is determined by how successful that stance is when applied. Aggregation and abstraction are epistemic principles: whether we will gain traction over a problem by an abstraction move is not a matter of ontology, but of the reliability of the process in the context in which it is used.

Reusable ontologies are clusters of terms that correspond with these levels, stances, perspectives, etc, to be each individually fitted on a situation to gain traction. Good examples are Hobbs' scale, change, cause, plan, and goal in [150]. The notion of a physical object with a location, an orientation, and movement (cf. [153, 183]) is another convincing example. Reusable ontologies sometimes go by the name of *core ontologies*, if we only consider the set of terms, or *core theories*, if we also take into account the non-ontological rules positioning them as a knowledge source.

Different KBS choose different levels of granularity, and a KBS may even switch between levels of granularity in the process of generating hypotheses. This does not imply that the underlying *ontologies* are somehow defeasible, or contextual. It is not an argument for nonmonotonic inference techniques in terminological reasoning. Rules recommending an abstraction, or effecting a shift between abstraction levels, are not ontological: if the decision to abstract eventually introduces incoherency it is the abstraction move that should be questioned. Rules recommending an abstraction move are defeasible.

The decision to *equate* concepts from different terminologies because they concur often or always in a specific KBS setting is obviously equally non-ontological and defeasible, and creates serious maintenance problems.

2.3.2. Ontological Stratification

The discussion of granularity in knowledge engineering relates to a central issue in philosophical ontology: the classification of qualitatively different types of existence (cf. for instance [254]) that will be referred to in this book as *ontological stratification* (following i.a. [55]). Strata are usually envisioned as deposited "on top of" each other, and objects in different strata are "co-located" against the same background, often spatio-temporal, with an intentional relation between things in one stratum to things in another one, and each stratum has separate criteria for existence in that stratum. Things in different strata explain each other in some way, but they do not *reduce* to each other.

Even on the physical level one can distinguish different strata that clearly do not reduce to each other in a meaningful way (cf. generally [84], and [254] for a discussion of phenomenological, microbiological, and physical-chemical apples). The distinction between physical, mental, social, and abstract objects is a typical example of stratification, as are Dennett's physical, design, and intentional stances, and the distinction

between ontology and epistemology made here. A well-written inquiry into the nature of legal reality is for instance found in [204]. A more general discussion of the rationale for stratification is however beyond the scope of this book.

The existence of the physical stratum can be demonstrated through the senses, directly or only in carefully managed experiments using special tools. The existence of the mental is presumably based on introspection. The abstract is created by a reasoner and identified by its purpose in reasoning. The existence of the social is grounded in (one's awareness of) recognition by multiple consciousnesses¹⁰.

There is a strong temptation, maybe grounded in Ockhamist reductionism (as [254] believes), and maybe simply in pre-existing epistemic habits¹¹, to assume that things that are co-located but inhabit different strata are in effect one and the same thing, conceived from different points of view. Obviously, we are part of reality, perceive reality through our experiences, and our experiences are contingent upon our ontological stance towards reality. Coincidence is however different from identity, even if it is often possible to fudge the consequences of the differences.

The *ontological distinction principle* (cf. [55]) prescribes that terms whose meaning is determined by different identity criteria must be disjoint, but we do not have to subscribe to this principle. The following weaker principle does hold without reservation:

Proposition 1. Ontological stratification principle: two concepts that describe objects in different strata do not stand in a terminological entailment relation to each other.

The road (physical thing) and the line (abstract thing) are not necessarily to be treated as disjoint terms: if the KBS only needs road-as-a-line objects, the road and its abstraction *can* be rolled into one, for the sake of simplicity, as long as we accept that predictions based on the line metaphor are correct for the road only within a certain margin of error. Instead of for instance keeping track of the fact that i' is an abstraction of i , and i a reference to some relevant thing in the outside world, one simply assumes $i = i'$, i.e. we make them co-referents on a logical level (even if one would, if pressed, nevertheless insist that they are two different things). Lines, surfaces, and volumes however cannot be equated without introducing inconsistency.

There are many instances of this coincidence issue. Well known is for instance the “Cicero = Tully” (or “Superman = Clark Kent”) identity problem (cf. for instance [224]). The agent identified by both Superman and Clark Kent for instance has the ability to fly, but only the social *role* of Superman is associated with flying. Do we need to distinguish the different behavioural expectations we have about the social roles of Superman and Clark Kent? Do we for instance dare to equate “the murder suspect” and “the murderer” in our minds, or are we open to the possibility that “the

¹⁰[198] for instance distinguishes between social objects in the narrow and broad sense, and gives quark and triangle as examples of ‘social’ in the broad sense because their meaning depends on communities that recognize them. Note however that it is not at all *necessary* for the community to exist for triangle to be a useful concept: it is a perfectly valid and useful abstraction for an individual consciousness. Abstract and social are distinct strata.

¹¹perhaps the same ones that lead us astray through metonymy and synecdoche (see section 2.3).

murder suspect” and “the murderer” may turn out to be different people? We know we have to separate them as soon as we encounter ontological incoherence, or bizarre entailments, but in many cases we can equate them without problems.

Chapter 4 will address the issue of why the identity of *institutional* facts and the *constitutive* facts they are grounded in should be conceptually separated, and why action and the task that is performed by the action should be separated. In chapter 5 the issue of why a document as a work, as an expression, as a manifestation, and as a (physical) item should be identified separately, as advocated by [245].

A KBS may hold incompatible conceptualizations of “the same thing” if properly stratified. It is however possible to organize knowledge components well, by clearly separating terms and terminologies on the ontology level into separate strata, without burdening KBS developers with an explosion of spurious abstract entities in the deployed KBS if this is unnecessary. In this book I will generally explicitly separate things for the purpose of clarity wherever needed, without subscribing to the principle that the domain and range of a relationship spanning ontological strata *must* be disjoint.

2.4. Epistemology

The epistemological dimension of knowledge representation accounts for the processes by which we acquire knowledge, and more specifically, how the knowledge representation realizes epistemic competence, i.e. the competence to solve a type of problem, in the hands of a rational agent that understands the knowledge representation. Central to this account is the notion of problem solving.

A problem is an obstacle which makes it difficult to achieve a desired objective, or for some machine to serve its purpose. In a broad sense, a problem exists when an individual becomes aware of a significant difference between what actually is the case and what is intended and expected: to solve the problem he has to bring about a change. The concept of *problem solving* is however rarely explicitly connected to the notion of a problem.

Standard descriptions of intelligent agents posit some three-step think-act-perceive cycle such as in [240]. This simple problem solving cycle, as sketched in figure 2.2, is at the center of the activity of problem solving: intelligent activity consists of monitoring, and comparison of observations to intentions and expectations, finding and understanding obstacles, or problems, characterized as significant deviations from the intended scenario, and planning, or designing, to remove them. The cycle positions two fundamentally different epistemic roles of knowledge:

(know-why) modeling and explaining the situation (cf. for instance [82, 258]), and **(know-how)** deciding what to do given the (explanation of the) situation (in accordance with Newell’s conception of knowledge in [209]).

This distinction is similar to the distinction between construction (how) and classification (why), or synthesis (how) and analysis (why) (cf. generally [59]).

Presumably we have hit upon a problem every time when deciding what to do next is non-trivial. Very often explaining the situation and deciding what to do can

be considered separate phases in problem solving, although one could argue that the optimal course of action should take into account all possible conceptualizations of the situation.

The three-step cycle loosely fits well on Breuker’s problem types in [59]. The major difference between this cycle and Breuker’s original acyclic proposal is the explicitness of the re-entry from remedy to planning and repair to design. In the case of repair, we can distinguish between repairing the machine in the case of malfunction of the machine relative to the specification, and revising the design if monitoring a simulation, scale model, or prototype shows it does not live up to expectations. Note that we do not only think in order to “act”. Lack of satisfaction with one’s diagnosis results will for instance generally result in a plan to monitor specific behaviours or even to coerce the occurrence of behaviours that help distinguish between alternative diagnostic hypotheses: in this case we act in the service of refining our explanation of the situation. The most trivial attempt to test a hypothesis is to *ask the user*: many KBS do nothing more than conducting a structured dialog with the user.

Note that most actual KBS do not support a single iteration of the whole cycle, but focus on aspects of it that are better – in the sense of effectiveness or efficiency – performed by a computer than by a human.

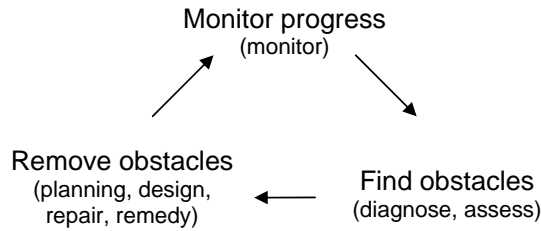


Figure 2.2. Problem solving cycle.

Considering the relation between the notion of “processes by which we acquire knowledge” and the notion of epistemic competence, we hit upon an ontological mismatch. Competence, whether physical, epistemic, or legal, naturally relates to *action*, following a decision, while an *epistemic process* suggests something that happens unintentionally, beyond conscious control. Apparently, we can distinguish between a kind of reasoning that “happens in the background”, beyond conscious control, and another part that is consciously and intentionally used or not. We at least have two different vocabularies to apply to reasoning.

In problem solving (method) literature we find both characterizations in terms of function – suggesting an embodied process, a machine, or a design specification for one, that can be used for some purpose – and in terms of competence – more suggestive of decision making and planning. Also to problem solving methods we can apply Dennett’s design and intentional stance of [95].

Terminological inference on the ontology is the most obvious candidate as a background process, since we already committed to ontology being not falsifiable within

a KBS, perhaps with other such processes we have not explicitly identified, while the interesting part of problem solving involves decision making on the epistemic level.

To solve a problem is to set oneself a task, and one has the epistemic competence to solve the problem if one knows an applicable (problem solving) method for performing the task. The problem solving method is characterized by a set of assumptions about the knowledge domain and the available domain knowledge that function as conditions for its application, a functional specification, and an operational specification (cf. [109]). The functional specification tells us what epistemic competence will be realized, and the operational one how it will be realized.

The terms describing the problem solving method belong to epistemic vocabulary: they exist to describe processes through which one reliably creates knowledge. Assumptions about available knowledge are *autoepistemic*, i.e. they are about epistemic entities, since they involve introspection over what one knows (or knows a reliable acquisition process for).

Sources of law obviously do not specify problem solving methods; At best they sometimes give instructions about the application of rules that presume certain things about the use of rules in the mind of the reasoner, and about available knowledge. To position these instructions, some explanation of my perspective on problem solving, and in particular defeasibility, is needed.

2.4.1. Ampliative Reasoning and Entrenchment

Problem solving methods generally are said to involve *abduction*, or inference to the best explanation. Inference to the best explanation is a basic method of reasoning in which one *chooses*, from a space of *hypotheses*, the hypothesis that would, *if* true, best explain the relevant evidence. Abductive reasoning is an account of *ampliative reasoning*: it adds information that was not already implicitly available in the premises. It is often described as a dual of deduction, as abduction allows the precondition ϕ of $\phi \models \psi$ to be inferred from the consequence ψ , and it is thus clearly non-deductive. But not every inversion of a deduction is obviously also a meaningful abduction; Making abductions is an epistemic decision making process, and one needs a good plan for that. Good strategy depends on what problem one is trying to solve.

The generic problem solving method par excellence, Generate & Test, is considered a generic abductive framework; Most problem solving methods in literature (cf. [76]) postulate additional epistemic assumptions or knowledge sources for generating hypotheses in the best possible order or cut down the portion of the search space that needs to be explored.

In addition there are many generic theories of default or nonmonotonic reasoning that can be considered accounts of ampliative reasoning. A shortcoming of these generic approaches is that they do not account for how ampliative inferences help to solve a problem, as opposed to problem solving method literature which does make this connection.

Parametric design is an example of a problem type that has a well-studied set of ampliative problem solving methods associated to it (see for instance [75, 207, 281, 280, 108]). Read these if the idea of a problem solving method remains unclear.

An abduction problem can be characterized as a tuple of sets of propositions¹² in conjunction (T, A, S) , where T is some background theory, S the situation to explain, and A a set of abducibles. A subset $E \subseteq A$ is an (abductive) explanation of S in a certain knowledge state if we believe that:

1. $T \cup E \models S$
2. $T \cup S \not\models \neg E$
3. $T \cup E \cup S$ is consistent¹³

The function of the abducibles is in essence to specify a set of propositions that may be assumed instead of (deductively) proved. We are willing to retract abducibles if they do not fit reality. There is a nice proverb about the role of these *assumptions*: “the wish is the father of the thought”. We do not make an assumption just because we believe it is true, but because it believing it *brings us closer to solving our problem*. The optimal strategy for explaining a situation therefore depends on the problem setting.

E is minimal if no proper subset E' of E is an explanation of S . One simple way to characterize *the best explanation* is as the disjunction of all minimal explanations. Typically one however wants to be able to explicitly choose between the members of this set of minimal explanations, i.e. to order them by (epistemic) preference, if explanations may be falsified at a later point in time by additional information.

Abduction problems can be conceived of in deductive frameworks in terms of model generation or satisfiability. The concept of abduction has been related in literature to various approaches to *defeasibility* such as belief revision and truth maintenance systems, default and other nonmonotonic logics, and to negation as failure in logic programming (cf. for instance [214] for an overview).

The *hypothesis* is an epistemic entity, a provisional idea that can be falsified by a test against evidence. A hypothesis has greater *explanatory power* if it provides greater opportunity for its own falsification. In a knowledge engineering setting, the objective is in my view generally to generate concrete plans, designs, arguments, or explanations, that:

1. are ontologically coherent, i.e. are not inconsistent with, are a valid model of, the terminological axioms;
2. pass some adequacy test, meet constraints and requirements, achieve some objective, etc;
3. can be executed, used, built, etc; generally, are at the right level of detail to be acted on; and
4. can be monitored by comparing intended or expected events against observable events to determine whether one encounters a new problem.

The characterization of abduction problems helps as a general framework for understanding how the first two objectives are met. The last two criteria are generally filled in implicitly in KBS, and refer to the required explanatory power. The KBS works

¹²I intend proposition in a general sense here: whatever it is that is expressed by a sentence that can be evaluated to true or false.

¹³Note that A need not be internally consistent, and it is not a given that E is.

towards a level of granularity and detail that is deemed suitable by the designers of the KBS. The simplest method of achieving this is to ensure that the reasoner has no more knowledge available than the knowledge required to produce an explanation on the right level of granularity and detail: by ensuring that reasoner has no knowledge that is irrelevant to the problem at hand, it is ensured that the reasoner does not perform irrelevant inferences. The set of abducibles is therefore usually carefully handcrafted to result in explanations on exactly the right level of detail.

The characterization of abduction problems given here fits better with the role of knowledge as explaining the situation than the role of knowledge as a resource for deciding what to do next: this is a result of referring to S as a situation, instead of, for instance, a plan or a design. To explain how a chess board can contain eight queens that do not attack each other is to solve the problem of placing the eight queens. S is sometimes conceived of as a goal, but this is merely confusing.

There is however a fundamental difference between explaining the situation and deciding what to do next, and it is most easily described in terms of decision making, of choosing between alternatives. Choice between alternatives reveals – from the external point of view – or is motivated by – from the internal point of view – a *preference* ordering. There are two fundamentally different ordering principles for generating and testing hypothesis:

1. The first ordering principle, modeled by *epistemic entrenchment* (cf. [117, 1, 208]), orders hypotheses on epistemic quality, i.e. the likelihood that they retain descriptive fidelity, or the order in which we adopt and abandon hypotheses to come to a solution. This is a preference relation on the *epistemic level*, which finds its rationale in an assessment of how reliable the process is by which it the hypotheses were generated. This ordering principle is the one modeled by defeasibility and nonmonotonicity in logic and knowledge representation.
2. The second ordering principle only comes into play when we consider reasoning directed towards a decision to do something: it models the order in which we adopt and abandon plans and designs. This is a preference on the *intentional level*, a utility measure on plans, designs, etc. In analogy with epistemic entrenchment, we may call this (with a neologism) *intention entrenchment*.

Example 1. I for instance prefer winning a lottery over losing a lottery, but I prefer *believing* that I will lose a lottery over believing that I will win a lottery, since this is likely to save me money in the long run. Too strong a preference for favourable outcomes will lead to unrealistic plans and designs.

The two different methods to order hypotheses find their origin in two different decision making processes. Epistemic level decision making – dominant in prediction, postdiction, diagnosis, and assessment – deals with explaining and evaluating what is happening, or *what to think*. Intentional level decision making – dominant in design, planning, and scheduling – deals with *what to do* and is normally speaking relative to the problem we are solving.

The distinction also plays a prominent role in classifying the epistemic role of rules found in sources of law. The two ordering principles are relevant to a number of legal theoretical concepts discussed in the next chapters. The norm is for instance based in

deontic choice. This appeals to the intentional level. Deontic logics also vary in the assumptions they make about this ordering relation in decision making. A notion like applicability on the other hand is to be understood in the context of epistemic-level decision-making.

Choice between alternatives and the composition of preference relations therefore plays an important role in the following chapters, both on the intentional and on the epistemic level, because of its central importance in explaining reasoning. Preference relations do not, however, very often explicitly play a role in KBS.

Managing both orderings in problem solving at the same time is a complex task. The best known even-handed approach for it is the *expected utility hypothesis* central to decision theory (cf. generally [137, 240]): if certain (not very plausible) assumptions are met, then utility and probability of alternatives can be composed into a single expected utility function that orders alternatives and can be used to choose between them. In knowledge engineering, expected utility and probability theory in general play a minor role. Expected utility is a “useful abstraction” in the sense discussed in section 2.3.1, but also one that has little cognitive plausibility and comes with very clear limits in applicability.

Since knowledge about generating hypotheses and knowledge about testing hypotheses are in practical KBS settings rarely in balance, both phases generally occur: even if we can compare any two alternatives to determine which one is preferred, we are not necessarily able to generate them in that order. We may not even be able to generate hypotheses consistent with the background theory: many practical approaches to abductive reasoning require filtering of the generated hypotheses. This creates a lot of variation in the knowledge we may have about the preference relations, and a lot of variation in approaches in knowledge engineering for generating and testing hypotheses.

The assumed mathematical properties of the preference order relations employed have a big influence on the efficiency of problem solving. It for instance makes a difference whether we recognize an adequate score immediately, i.e. we are dealing with a *scale* and can safely discard the hypotheses not yet seen because we have hit a sufficiently high or the highest score on the scale. If we can generate hypotheses in the correct order, then the test is reduced to confirming, for the KBS in general, that no better hypothesis exists than the one generated. This is a good reason to make assumptions, for instance the ones underlying expected utility, even if one knows they are occasionally falsified in practice.

The notion of a preference ordering to explain choice is not entirely straightforward: the hard problem of knowledge representation – and of decision making – is generally to enumerate the sensible alternatives. As pointed out by for instance Sagoff in [241] preference as revealed by choices appears empirically well-founded in the notion of choice, but the choice problem is in the end as unobservable as the preference. The following joke explains the issue well:

Two graduate students overtook Professor Paul Samuelson as he walked. “There’s a beggar at the corner,” they told him, “who, when offered the choice between fifty cents and a dollar, always takes the fifty cents.” Samuelson replied, “He’s irrational.” Samuelson decided to see for himself. “In my left hand, I have fifty cents; in my right, one

dollar; you may have whichever one you prefer,” Samuelson said to the beggar. “I’ll take the fifty cents,” the man answered without hesitation.

After giving him the two quarters, Samuelson asked, “Don’t you understand that a dollar is worth twice as much as fifty cents?” “Of course I do.” “Then why did you take the fifty cents?” “Had I taken the dollar,” the beggar replied, “economists wouldn’t troop down here every day to offer me the choice.”

The set of alternatives ascribed to the beggar and chosen from by the professor and the set of alternatives chosen from by the beggar are not the same set: the professor failed to generate the right explanation of the behaviour of the beggar.

Both proposed preference orders suffer from serious cognitive plausibility problems, also recognized in law. As for instance pointed out by Mommers in [202] the idea of the autonomous court of law presupposes *doxastic voluntarism*, i.e. the stance that one can consciously choose whether or not to believe something. Amsterdam and Bruner in [6] attack this very proposition, based on evidence from psychological experiments, but the legal system cannot function without this presumption. The same may obviously be said of behaviour compliant with norms: the legislator and the court of law must assume that human beings are capable of consciously integrating the law into their behaviour, and it must ascribe choices between a concrete set of alternatives to them.

Law cannot function without the presumption that the sources of law can influence choices between alternatives, and that these choices are observable from the outside.

Preference relations also have the redeeming feature that they characterize underlying problem solving strategies in a simple way, and link them directly to available logics and algorithms, regardless of whether they are cognitively very plausible descriptions of intelligent behaviour.

In chapters 4 and 6 I often explain how the law’s commands are integrated into behaviour in terms of choices and preferences, for no better reason than that legal theory also tends to defer the issue to *rational choice theory* (cf. Hansson in [137] for an overview of choice theory).

2.4.2. Epistemic Things and Epistemic Roles

Problem solving strategy is conceived of in terms of processes and decision problems involving epistemic things. Epistemic things are mental things: a mind “has” them, and the representation refers to them. If the “mind” is the KBS’s, mental and abstract collapse into one for practical purposes; Let’s call this the *internal* perspective. Since the mind makes its own epistemic things, it may in principle also be able to maintain perfect fidelity. For the KBS, a proposition and its representation in the form of a data structure are one and the same thing. This is not the case with hypotheses about what somebody else knows; This is the *external* perspective: the epistemic object is *ascribed* to someone else. The KBS may need a theory of both, and it does not necessarily use the same theory for the internal mental world and external mental worlds.

Of particular importance to knowledge engineering are the epistemic entities I will refer to as *doxastic*¹⁴ or propositional attitudes: these express mental attitudes towards, or the epistemic role or status of, propositions. These form the metaphorical

¹⁴i.e. pertaining to beliefs.

bridge between our conceptualization of the mental world and the knowledge representation language that expresses it.

Searle’s speech act theory (cf. [248]), one of the dominant perspectives on language, is founded in a strict distinction between the proposition expressed by a speech act, and the speech act’s meaning as an act, grounded in a theory of action. We can make a similarly strict distinction between propositions and the epistemic objects that contain them. Problem solving methods, the abductive framework sketched before, and logics in general, can then be conceptualized as theories of epistemic process or action.

Logical entailment $\phi \models \psi$ is a relation on propositions, independent of their epistemic role or status as a belief. The *belief*, a simple proposition container, is the fundamental building block of the mind from a doxastic perspective. Commitment to the existence of individual beliefs with propositional content justifies reification and manipulation of the sentences of the knowledge representation language as if they were beliefs. This allows me to make epistemic level statements such as “believing ϕ is better than believing ψ ” or “ ϕ is a terminological axiom”. An account of reasoning is a account of sets of beliefs held in a progression of knowledge states.

Epistemic terms are found everywhere: hypotheses, factors, evidence, symptoms, indicators, problems, solutions, etc (cf. [26]). As pointed out before (see 2.3, about metonymy), epistemic vocabulary should not be confused with the domain ontology: flue is not a “subtype” of symptom in the medical domain, but often has the *epistemic role* of symptom in medical diagnosis.

Epistemic vocabulary often depends on whether we take an *ex ante* or *ex post* perspective: we hypothesize about actions or events in the past or future, or we reason about designs and plans before or after we start executing or using them. *Ex ante prediction* and *ex post postdiction*, or, often, simply, explanation, are basic abduction problems, and do not connect to action.

Breuker’s problem types in [59] can also be classified on this dimension: planning, design, modeling, parametric design, and scheduling are conceptualized from an *ex ante* perspective, while monitoring, diagnosis, and assessment are conceptualized *ex post*. Remedies and repairs should be understood as re-entries into planning, but described from an *ex post* perspective. Breuker’s problem types are conceptualized in the context of decision making and action.

One of the complicating factors of law, is that it also explicitly uses epistemic terms with complicated semantics, to be used from both internal and external perspectives. We need to explain the LKBS’s operations in terms of them because proper justification, with reference to the sources of law, is a central issue in law. Accounts of reasoning given in legislation and legal theory often lead us away from the approach that would be chosen for the same type of reasoning problem in other knowledge engineering fields.

Take as an example for instance the doctrine of *mens rea*, the guilty mind, which distinguishes between intention (direct or oblique), knowledge, recklessness, and negligence¹⁵: all of these are epistemic modalities, to be ascribed to a suspect, and together they form a simple theory of the functioning of the mind. These concepts will however

¹⁵Or the comparable Dutch distinction between oogmerk, opzet, voorwaardelijk opzet.

not be found in automated planning literature. As I will explain in chapter 4 theories about norms in law are also strongly influenced by the distinction between the *ex ante* and *ex post* point of view.

Summarizing, it makes a difference whether we are taking an internal or external perspective to epistemology. The internal perspective is the one we take to describe automated reasoning by the KBS. The law takes an external perspective to epistemology when it ascribes intent or knowledge to agents. The way we describe reasoning depends on our position relative to the thing we are reasoning about (*ex ante*, *ex post*), whether we are dealing with a machine or plan, and whether we are making a decision or simply trying to understand something. In any of these contexts, we can make more and less plausible inferences, and when we are making a decision we choose between better and worse alternatives.

Finally we get to the relationship between epistemic objects and the things they are about. It is very common to simply equate the things we are reasoning about with the variables and constants of our logical language, while epistemic roles characterize the sentences we can create in the language: epistemic role is a qualification of a reified proposition. The language used in this book will be introduced in section 3.2. This brings us to the next level: the level of language and information. By making knowledge explicit, we turn it into information.

2.5. Argumentation and Information

More often than not, knowledge based systems do not reason in order to decide what to “do”, but simply tell the conclusion of reasoning, and its explanation and justification, to the user. In [59] the explanation and justification are even considered essential parts of a solution.

Explanation and *justification* are two quite different things. When you explain something, you do not have to take a position on whether it is good or bad. When you justify something, you try to explain why something is good. They belong in different vocabularies – one for problem solving and one for decision making. Generally speaking one solves a problem, and explains the solution, as opposed to making a decision in a situation, and justifying the decision. Decision making involves a preference ordering of alternative outcomes, and generally involves a design or plan, while explanation applies to any kind of problem solving and generally focuses on epistemic entrenchment; Hence the earlier characterization of abductive reasoning as “inference to the best *explanation*”.

Since, as noted in the previous section, epistemic entrenchment can also be conceived of as rational decision making on the epistemic level, it is possible to recast explanation into decision making terms. One can for instance *explain* how one arrived at a taxable income, or *justify* one’s income tax claim.

Explanations and justifications are constructed from arguments. An argument is typically a *claim*, or “thesis”, expressing a proposition, backed up with evidence that supports the claim.

Argumentation is primarily an *information* issue, and not an ontological or epistemological one, although there are obvious connections between good problem solving

strategy and good argumentation.

If one engages in discourse one exchanges *information*, in the form of propositions. By doing so, one commits oneself to these propositions; they become part of a shared discourse context, i.e. the parties in the discourse, including yourself, know you have claimed these propositions are true and they expect that you also believe that propositions they believe are entailed by it – according to them – are true (cf. generally [227]). The same logical entailment that operates on sets of beliefs can also be applied to this shared discourse context, but the discourse context has no direct relation to belief sets: it is perfectly possible to commit oneself to propositions one does not believe are true, and the totality of exchanged propositions obviously need not be consistent.

Information is a message received and understood. Information is a quality of a message from a sender to one or more receivers. Information is always about something: it is the proposition represented by an act of representation. It assumes the existence of a common language understood by the sender and at least one of the receivers: it is that which would be communicated by a message if it were sent from a sender to a receiver capable of understanding the message, i.e. that knows what the message is about.

Since the existence of a definite sender is required, information cannot be directly extracted from an environment, e.g., through observation or measurement. Knowledge on the other hand can be. Knowledge and information are different things. Observations and measurements do *become* information by representing them and “sending” them to some receiver capable of understanding them.

The study of information exchange is different from the ontology and epistemological considerations that determine what the information is about. As pointed out in section 2.3, information exchange is ruled by principles (such as for instance the Gricean maxims of quality, quantity, relation, and manner; cf. generally [227]) and tactical considerations (cf. for instance [230]) that are foreign to ontology and epistemology. Presentation of argumentation in court, or any other setting for that matter, is not a direct representation of the underlying legal reasoning, but a competitive game. It is in fact a separate planning problem for the involved parties (as for instance explicitly pointed out in [256]).

The KBS can be conceived of as a machine that sends and receives, in accordance with a predefined plan, information conformant to a predefined schema. In typical KBS use cases a human provides a formalized problem description to a computer, and receives a solution and explanation to interpret. It is however good to realize that the roles in the transaction can also be reversed: the computer asks a person or a large number of people to solve a problem, then collects, interprets, and integrates their partial solutions. This idea – *human-based computation* – aims to optimize the use of differences in abilities between humans and computer agents, and is already used for image and video rating and labeling on the World Wide Web on a voluntary basis. It is obviously not always the KBS that plays the role of knowledge source: KBS will increasingly collect knowledge from “the wisdom of the crowd”.

If a KBS engages in a structured dialog, i.e. it doesn’t just answer questions, but also poses them, it implicitly creates a division of labour between itself and its

human user(s). Implicitly this division of labour either distributes a burden of proof, or embodies assumptions about ‘who knows what’.

The plan that determines the structure of the dialog between KBS and user is very often a derivative product of problem solving strategy or the logic used. In some cases ([57] comes to mind) the argumentation, or explanation, for a certain thesis is explicitly equated with the *proof tree* produced by the problem solving process. For instance [272] describes our own approach to compiling proof trees out of general rule sets in an application intended to help clients explain a request for legal assistance in legal terms.

Although there are additional considerations to take into account for good argumentation, it is obvious that explicit argumentation must closely follow the problem solving strategy; Firstly, the argument made in section 2.2.2 would be meaningless if no such close relationship existed, and secondly the audience, presumably equipped with similar symbol processing machinery, is supposed to *understand* the message. Obviously, *doxastic* epistemological theories are in a sense always based on the presumption that argumentation reveals the working of the underlying epistemic process.

The abductive problem setting sketched in the previous section for instance clearly suggests an interaction form: if possible explanations are sets of abducibles, it makes sense to translate the abducibles into questions to the KBS user in the order the KBS runs into them. A slightly more refined approach asks for the most discriminating abducible first, i.e. an abducible that cuts our set of possible explanations into to equal halves. See for instance [259] for a useful general theory of asking relevant questions.

While the information and argumentation level are very interesting to discuss in relation to knowledge and reasoning, we need to keep in mind that knowledge and information, and reasoning and argumentation, are indeed different things. When one implements a KBS one does need to consider the interaction between KBS and users, but the reasoning process does not need to directly mirror its explanation, its justification, or any kind of related argumentation encountered in the field. In this book both the human computer interface of the KBS and legal argumentation play almost no role in the explanation of knowledge representation decisions. The perspective taken is strictly that of one *mind* solving legal problems.

2.5.1. Adversarial Dialogs

A mode of message passing that deserves special attention is the dialectical or adversarial one, often referred to as (legal) argumentation in the narrow sense, where the participants exchange arguments and counter-arguments for a thesis or antithesis. These argumentation games are often ruled by a procedure and assigned burdens of proof and information duties. This is most apparent in court cases, which function as a more or less level playing field for argumentation. Many other legal exchanges of claims have a similar form, but with one dominant side who makes the rules and for instance a suppliant who claims entitlement to a benefit.

Because the court is the prototypical setting of the legal system, it has a large influence on the way legal problem solving is conceptualized in legal theory. In

computational legal theory we find a large school of thought that places adversarial argumentation on the center stage (cf. generally i.a. [278, 122, 230]). Defeasibility and other such violent metaphors characterize this approach. When one models an adversarial dialog *setting*, i.e the discourse context instead of the content of the minds of the participants, then some of the observations made earlier no longer apply: since the participants in argumentation are not committed to agreeing with each other on any subject, even ontology can become a subject of discussion and is therefore *defeasible*, etc.

Since case law is obviously one of the most important knowledge sources for LKBS, there is an argument to be made for modeling legal argumentation, the procedures that guide it, and the “good moves” that are made as we find them, instead of trying to distill coherent epistemological “deep structures” (cf. [93]) underlying them. This is a point explicitly made for instance by Gordon in [35]. Since issues of ontology and epistemology are rarely explicitly addressed in case law, they appear relatively unimportant.

The relative rarity of arguments about ontology can however also be interpreted in favour of the importance of ontology and epistemology in knowledge engineering as I pointed out in [30]. Deconstructivist Schlag explains in [247] that the games of law can be played (and won) even if one doesn't know what the ball (the law) one plays with looks like: no one engaged in “doing law” will raise the question. Superficially this supports the view that ontology is unimportant. Schlag (cf. [247]) proposes that the following rhetorical hierarchy is what actually guides legal practice:

1. Do not confront an ontological question if it can be handled as an epistemic question.
2. Do not confront an epistemic question if it can be handled as a normative question.
3. Do not confront a normative question if it can be handled as a technical question.

Ontological questions question the truth of terminological axioms, and the ontological inferences based on them. Based on the examples given in [247], it is clear that epistemic simply means (indirectly) questioning (the reliability of) a non-terminological inference made, from certain premises that function as evidence and that are not themselves questioned, to arrive at a thesis. Normative questions address whether something is allowed or disallowed, good or bad, etc. In each of these cases the inference is in law generally based on a rule, and it is either the existence of the rule, the accuracy of the description of the rule, or the applicability of the rule to the case under scrutiny that is questioned.

Technical simply appears to mean “relating to the (truth of the) facts of the case”. It concerns questioning those propositions, not (yet) established by an explicit argument from premises, whose truth is left to common sense. These propositions do not generally have the character of a rule.

This rhetorical hierarchy is relevant to legal knowledge engineering because because it exposes a theory about *epistemic entrenchment* that normally remains implicit. The hierarchy tells us which propositions to put forward and which (parts) of the opponent's propositions to attack. It tells us that explicitly attacking terminological

axioms is the weakest move one could make, in court and, I would propose, outside it. Questioning ontology and the quality of epistemic processes directly attacks the very notion that a shared discourse context exists and that we can perform logical entailment on it. Moving to these levels, certainly the ontological one, is bound to provoke irritation.

This hierarchy has practical application in legal knowledge engineering. The following list tentatively orders the four levels mentioned by Schlag by decreasing reusability in law, and therefore by decreasing usefulness as a *reusable* knowledge component:

1. Ontology
2. Epistemology
3. Normativity
4. Technique

Ontology is highly resistant to change, and is assumed to be *not* falsifiable at all in KBS. The rhetorical hierarchy corroborates that design choice. Even ontology is in the end subject to questioning, at least in its sense of being a shared understanding between discourse participants, but certainly not in a routine decision making context. Falsification of an ontological axiom is rather a reason for knowledge engineers to review the decision problem solved by the KBS altogether.

Epistemology and normativity can be interpreted as relating to the what to think (epistemology) and what to do (normativity) distinctions of section 2.4.1. Schlag makes a clear choice here, by making rules guiding interpretation of a situation more resistant to being questioned than rules guiding decision making. Also this is in line with section 2.4.1: interpretation of the situation (or finding obstacles in terms of section 2.4) is indeed logically prior to taking action in it (removing obstacles).

In practical KBS both are often not falsifiable: the KBS, whose knowledge base is expressed in a knowledge representation language that exhibits the property of monotonicity, reflects one single unambiguous interpretation of the knowledge domain relative to a specific problem. In AI & Law theory one however also finds many proposals that consider all legal reasoning falsifiable. An intermediate form also exists; In for instance the framework of [265] (but also my own work in [286]) the possibility of falsification of normative qualifications *is* taken into account, but situation recognition is assumed to be a mere matter of translating the case into the terms of the ontology. The possibility of falsifiability of the interpretation of the situation is not considered. One does not easily find proposals that are counter-examples to the hierarchy proposed by Schlag.

The defeasibility of arguments in a discourse with more than one participant is a fundamentally different issue from defeasibility of rules in reasoning, when reasoning is equated with problem solving. Ontology is not defeasible in problem solving. In argumentation it may be, but even the tactical rules for argumentation proposed by Schlag acknowledge that questioning ontology is a rare and often even ill-advised move.

2.6. Conclusions

Reasoning and knowledge as conceived in this book are based in the concept of reasoning as *problem solving*, and knowledge as a resource in problem solving. Knowledge directly relates to descriptive *fidelity* and to *problem solving competence*, and only indirectly – after one has committed to a specific doxastic theory of reasoning – to information and its structure.

In this book therefore both the human-computer interface of the KBS and legal argumentation play only a minor role in the explanation of knowledge representation decisions. The perspective taken is strictly that of one *mind* solving legal problems, and the knowledge ascribed to it to explain its apparent problem solving competence.

On the other hand we need to be aware that one cannot avoid assuming the existence of some interesting relationship between beliefs in reasoning and arguments in argumentation, and more generally, that information reveals something about the processes that created it and use it – even if some intelligent “decoding” is necessary. The use of a knowledge representation language of course presumes some commitments to an interpretation of the mental world in terms of structured information. In this book, we are committed to the description logic OWL DL.

The objective of this book is however hard to align with the notion of knowledge as problem solving: reusability of knowledge components that represent a source of law – not a problem or task – must be increased by keeping considerations of problem solving competence at arm’s length. Separation of reusable knowledge about the domain and reusable knowledge about problem solving, or ontology and epistemology, as characterized in [26, 153, 154] is the first step to designing reusable knowledge components.

For the subject of this book this means that we are interested in ontological knowledge about the world of law, *as expressed* in the sources of law. Epistemological knowledge plays a role to the extent that the sources of law *themselves* model that mental world.

The sources of law will however generally use epistemological concepts and constructions different from those used in knowledge engineering, while one of our goals is a general compatibility with general knowledge engineering work on problem types like planning, assessment, etc. In chapter 4 a survey of relevant legal concepts and patterns is made, with the aim of divorcing epistemological background – that often just explains the main functions of legal constructs from certain especially relevant points of view – from the ontological essentials that are represented. The resulting representation of *legal rules*, to be developed in chapter 5, is often very simple and straightforward.

Section 2.3 identified some information-level issues that complicate the design of ontologies: polysemy, metonymy, synecdoche, etc. When one tries to distill ontological knowledge from written text, these complications should also be kept in mind: one cannot sincerely take a written text “literally”.

Section 2.3 about ontology makes a distinction between concrete objects that are a surrogate for something out there (mental or physical), and abstract objects that only take their meaning from what you can do with them on an epistemological level.

Contrary to physical and mental concepts, abstract ones can usually be fully defined (with necessary and sufficient conditions). This distinction between a physical, mental, and abstract world is an example of *ontological stratification*, the classification of qualitatively different types of existence (cf. for instance [254, 55]).

Strata are usually envisioned as deposited “on top of” each other, and objects in different strata are “co-located” against the same background, with an intentional relation between things in one stratum to things in another one. This identification between structures in different strata, and the perception that they are layered on top of each other tells us something important about epistemology.

The common application of ontological strata deposits one or more abstract layers on top of a concrete one. The move from a concrete thing to an abstract one identified with it, is *abstraction*. Abstracting a concrete thing is done in order to use the associated abstract operational theory to explain something about the concrete thing. By doing so one temporarily makes the abstract thing a surrogate of the concrete thing. “Scientific style” theories recommend an abstraction to an abstract, mathematical thing: the inferences that can be made inside the theory after this abstraction is made are not considered *hypothetical*, but rather the rules telling us when to abstract¹⁶.

Section 2.4.1 made a distinction between non-ampliative, deductive reasoning, and ampliative, hypothetical reasoning. The realization of non-trivial problem solving competence invariably involves ampliative reasoning. A premise for the rest of this book, to be applied to the law, is that the direction in which we hypothesize involves a change of ontological stratum. Only for rules introducing new objects in another ontological stratum defeasibility plays a role. The conceptualization of an ontological stratum is by definition not defeasible within the problem solving process: instead one may switch between conceptualizations in order to “gain traction” over concrete reality.

If ampliative reasoning involves defeasibility, then there are alternative hypotheses (actions, situations, beliefs) to choose from and a preference relation orders the alternatives. Ampliative reasoning is therefore often characterized as a decision problem –involving reasoning actions– while non-ampliative reasoning is a process. Very often a KBS manages two interacting decision making problems at the same time: an epistemic level one, on beliefs, to address defeasibility, and an intention-level one, on actions, or the situations one expects as a result of actions. Knowledge of these preference relations again will play a role only to the extent that the sources of law *themselves* model it: no attempt will be made to model cognitively plausible or logically convenient choice principles that make problems solvable. Chapter 3 provides us with some basic instruments for defeasibility. In chapter 6 I will have to appeal to the intention-level preference ordering to explain *normative order*.

The following chapters will attempt to stratify law. Special attention will go in chapter 4 to intention, a mental thing, and institutional things, which are social in nature.

When one is representing (knowledge about) sources of law, *isomorphism* between

¹⁶E.g. apply $F = ma$ if you can determine a candidate m and a , but only if it is safe to assume that m remains constant. Consider for instance the application of the formula to a snow ball thrown from a building or rolling from a slope.

the source and its representation is of course of great importance. It is an explicitly stated objective of this book. Isomorphism in the sense used in section 2.2 means that there is some structural coincidence between the source of law and its representation.

This structural coincidence is in this book however not based on identifying text fragments in the source of law with logical propositions in the knowledge representation language. Makinson and Dubislav (in section 2.2.2) rightly criticize this identification for norms, but this argument in my view also extends to legal rules, or any other things “contained” by the source of law. Instead there is a structure of objects in the knowledge domain that represents the meaningful objects found in the source of law, and the logical propositions that we qualify as knowledge are *about* these objects. This also involves a move between ontological strata.

This approach to representation and isomorphism will be worked out in chapter 5.

Chapter 3

Knowledge Components

3.1. Introduction

Central to reusability in legal knowledge engineering is the notion of a knowledge component, in particular the knowledge component representing a source of law. Chapter 2 mainly focused on types of knowledge that should be differentiated in order to maximize the reusability of knowledge. For the knowledge component representing a source of law this means *not* including knowledge of generic knowledge domains, and *not* including a theory of legal reasoning that clearly aims at a specific problem solving competence and makes assumptions about generic problem domains for that purpose.

This chapter addresses the question what a knowledge component is in the context of the Semantic Web. Section 3.2 introduces the Semantic Web and its core technical standards, the uniform resource identifier and the Resource Description Framework, of which OWL DL – the knowledge representation language of interest in this book – is a component. Important to understand about the uniform resource identifier is that it does not meet the so-called *unique name assumption* often made in knowledge representation and information description logics. The *unique name assumption* will return in section 3.4.

Section 3.3 gives a classical description based on Levesque’s notion of knowledge sources in [187] of what a knowledge component is and how a KBS interacts with one. Of special interest is the possibility of *revision* of a knowledge base as a partial account of defeasible reasoning.

In section 3.4 the OWL DL logic is characterized. The design requirements for OWL DL are irreconcilable with defeasible reasoning, given the open nature of the Semantic Web. The OWL DL resource, as oppose to a knowledge base, is explicitly intended to be automatically fetched from the Internet, combined with other OWL resources if they are consistent, and automatically reasoned on. For defeasible logics such operations have unpredictable consequences, and the Semantic Web standardization body (of the World Wide Web Consortium) consistently chooses against features that would make logical inference defeasible on the Semantic Web.

In law we do find rules that are defeasible, and that must be accounted for as knowledge with an ampliative function (cf. section 2.4.1). As explained in section 2.4.1 ampliative reasoning is essentially hypothesis generation, and hypotheses are falsifiable.

This of course does not mean that one cannot account for any form of ampliative inference in a purely deductive framework. As pointed out in section 2.4.1 ampliative reasoning is often characterized as a dual of deductive reasoning. A sufficiently articulate deductive theory of some knowledge domain may be sufficient to approach ampliative reasoning as model generation or satisfiability. Section 3.4.3 gives the example of solving the eight queens problem on a chess board: the rules of chess restrict the domain sufficiently to address the problem of placing 8 queens on a chess board as systematically generating the 92 fully articulate models by brute force of the proposition *eight queens on a chess board that do not attack each other*. We can do this because we abstractly know the precise extension of the relevant chess terms.

In the absence of such articulate knowledge, and this is generally the case in law, as well as elsewhere, the ampliative knowledge we use for model expansion is not a

direct derivation of our conceptualization of the domain. In this case we use a strategy for navigating through an unknown space of possible models, and the rules describing this strategy are defeasible.

There are several possible perspectives (and of course the combinations of them) for epistemologically accounting for the meaning of such defeasible rules, for instance:

The autoepistemic perspective: The knowledge we have about a case – i.e. the propositions we have committed ourselves to – restrict the possible interpretations of a case, and these defeasible rules provide interesting hints about what other interesting knowledge is attainable¹ given what we know; In extension we can say that rules represent advice of the form *if we know [..], (we may) assume that [..]*;

The teleological perspective: The direction in which we want to expand our knowledge, either by making assumptions or asking questions, depends on the nature of the problem we are solving; In extension we can say that rules represent advice of the form *if we want to conclude that [..], assume that [..]*; and

The normality perspective: If we make assumptions, we prefer to assume that things are as normal as possible; the case is minimally abnormal.

In section 3.4.2 I will give an account of the first perspective, autoepistemic or introspective reasoning, interpreted as a form of interaction with the knowledge components of section 3.3. The second and third require too much epistemological commitment for the purposes of this work (cf. 2.6), although the realization of problem solving competence would not be possible without them. The second perspective is taken if one considers problem types and problem solving methods. The general epistemic framework of *minimality* (as Makinson calls it in [194]) and similar accounts of normality allow one to develop cognitively plausible and fully articulate logics of defeasibility that automatically and efficiently produce plausible and *minimal* interpretations of the case.

There are two reasons to account for the autoepistemic perspective in preference of the others:

1. We cannot get around the fact that what we know imposes at least some semantic constraint on the applicability of defeasible rules; and
2. The autoepistemic interpretation of interaction with OWL DL is used so widely in Semantic Web applications, apparently sometimes unknowingly, that we can take it for granted as a design commitment for KBS based on interaction with OWL DL knowledge components (cf. section 3.4.1).

It should however be clear that the discussion in section 3.4.2 aims to explain certain autoepistemic constraints on reasoning, and not a fully articulated reasoning method.

The subject of this book is representation of sources of law as knowledge components, while minimizing commitment to generic purpose theories aimed at realizing a certain problem solving competence. The theme that will be explored in chapter 4 is the relation between the law and intelligent – that is, knowledgeable – behaviour, with

¹Without however accounting in any way of why a certain assumption is interesting.

the aim of finding some minimal ontological commitments we have to make to ground the law in a common sense understanding of the world.

Section 3.5.1 in this chapter explores the notion of source of law. Directly linking knowledge components to sources of law as a knowledge source, presupposes that the sources of law *mean* something, *without* an appeal to a specific epistemic competence. Referring back to the earlier observation on *doxastic voluntarism* (cf. section 2.4.1), it seems that they do in as far as the interpretation of the sources of law is not *completely* arbitrary. On the other hand a knowledge component that represents a source of law does not exhibit a specific epistemic competence relative to a specific problem type: only a KBS deployed in a specific setting does.

The law, and legal theory, appeal to a shared common sense understanding of the world – including the mental world. Section 3.5.2 is a quick survey of candidate generic purpose knowledge sources – knowledge about actions, about granularity, about written documents as carriers of information – that are a prerequisite for explaining the way law works. Since there are many alternative accounts of these, there are necessarily also alternative accounts of their application in law, each of them too specific to function as a conceptual coatrack for explaining the purpose of knowledge of sources of law.

The primary purpose of this chapter is to set up a shared understanding of what is intended with a knowledge component, and how defeasible rules interact with ontology. This chapter will also introduce logical syntax that will be used throughout this book.

3.2. The Semantic Web

Most of the work this thesis was performed in the context of the *Semantic Web*. The Semantic Web, and IT standardization in general, are often proposed in political-administrative circles as *potentially* the most effective non-legal solution for reducing the *administrative burden*, as pointed out in section 2.1.1 and [7].

The World Wide Web Consortium² is responsible for the technical open standards that make the World Wide Web possible: HTML, XML³, HTTP, RDF, etc. The World Wide Web is mostly oriented towards information that is more or less directly presented to human beings: text, images, movies, etc. A major portion of Internet traffic however consists of data that does not fit this category: it is solely understood by computer programs. Both types of information may use for instance the XML standard, but there is a fundamental difference in meaning.

For most purposes there is no material difference between an XML structure retrievable from the Internet and the document, or movie, or image it describes. Only when one considers ephemeral issues such as how many copies of the document are around one starts noting a material difference between the familiar document and its simile on the computer. It has therefore become a convention that a XML data structure representing a document *is* a manifestation of that document, and not a description of it. For arbitrary XML data, for instance inventory records, exchanged between computer programs this convention however doesn't always work: they clearly *describe*

²<http://www.w3.org>

³<http://www.w3.org/XML>

something, for instance the inventory of a warehouse, but they clearly *are* not that thing.

The distinction between information as being understandable to humans and information understood only by computer programs eventually gave rise to a conceptual distinction between:

- The World Wide Web for human-understandable data; and the
- The Semantic Web for computer-understandable data.

World Wide Web standards have been designed to refer to an accessible world consisting of data structures that can (usually) be retrieved in some way over Internet protocols. Semantic Web standards often refer to an in principle inaccessible one consisting of things “out there” that can only be referred to by way of a symbol that functions as a substitute for the thing itself.

Of particular interest to the World Wide Web Consortium in its conception of the Semantic Web is *metadata*, often characterized as information about information, or data about data, and in particular computer-understandable data about human-understandable data. The primary use case for the Semantic Web is to make the World Wide Web better, more intelligent, by encoding computer-understandable information about documents. This is also the main subject of this book: computer-understandable information about sources of law.

Critics of the Semantic Web vision of standardizing metadata point out that the same computer-understandable information can be described in very many ways, that schemas are not viewpoint-neutral, and that people are generally stupid, lazy, and opportunistic liars (cf. for instance [98]). While most of these points have already been addressed here in some form, the point about opportunism is specific to the Semantic Web:

Metadata exists in a competitive world. Suppliers compete to sell their goods, cranks compete to convey their crackpot theories, artists compete for audience. Attention-spans and wallets may not be zero-sum, but they're damned close. [...] Meta-utopia is a world of reliable metadata. When poisoning the well confers benefits to the poisoners, the meta-waters get awfully toxic in short order. **In [98]**

This perhaps explains why the Semantic Web so far has not made a big impact on the general web⁴. Semantic Web standards have however been adopted quite enthusiastically by the knowledge engineering community for KBS development, with the understanding that the reliability of Semantic Web data retrieved from the Internet depends on the reliability of its source: a KBS does not use arbitrary metadata available on the Internet. Instead some fixed content syndication arrangement, involving a license fee, is expected, or imported knowledge components come from a trusted standardization body.

The World Wide Web uses the same identifiers – uniform resource identifiers⁵ (*URI*) – as the Semantic Web. The fundamental difference is that the URI in the World Wide Web is used to attach identifiers to XML data structures and other

⁴But note that for instance RSS (Really Simple Syndication, RDF Site Summary) is actually used extensively under the hood of many large web sites.

⁵<http://tools.ietf.org/html/rfc3986>

information objects in order for software to refer to and locate these objects, and in the Semantic web to refer to things in an often inaccessible domain of reference “out there” that are *described* by way of XML data structures (see figure 3.1), which in turn can be identified by URI and located on the World Wide Web.

The URI is obviously not a unique identifier: there is no bijective relation between the identifier and the thing it identifies.

The URI starts with a domain name, which has a *legal user* who has a leasing agreement with a domain name registrar. If users respects each other’s legal rights, and legal users use each URI only once to refer to something, then URI are globally unique identifiers. This is also one of the reasons why URI cannot possibly be bijective: the legal user of the URI would obtain a unique privilege relative to the object identified if no other URI could be made to identify that object.

3.2.1. The Resource Description Framework

The core Semantic Web standard is the Resource Description Framework⁶ (*RDF*). Compare the following two XML statements:

1. `<name xml:base='http://foo.org/foo' xml:id='bar'>Alexander</name>`
2. `<xmlNameElement rdf:about='http://foo.org/foo#bar'>
 <rdf:value>Alexander</rdf:value></xmlNameElement>`

The first one is an XML element defined in some XML schema with a URI identity marker `http://foo.org/foo#bar`. The URI identifies the XML element `<name>Alexander</name>`, and not its value “Alexander” or the fact that the value is a name. The second statement is RDF, and encodes a (rather trivial) statement *about* the thing referred to by the identity marker `http://foo.org/foo#bar`, being `<name>Alexander</name>`. This difference in interpretation cannot be inferred from the XML but is a difference in the semantics of basic XML (specifically `xml:base` and `xml:id`) and RDF.

The interface between retrievable XML and the outside world consists of these URI. The document as manifested in XML *is* what the URI refers to, while RDF *describes* the thing referred to by the URI. Obviously it is possible to describe everything, including the structure, contained text blocks, references, etc. in the document format in RDF.

In this book we use URIs as constants for terms, for individuals, for variables, for information items in general, for doxastic entities *qua* information items (as discussed in section 2.4), and for ontologies and knowledge components.

In place of (absolute) URI one may use URI references: these resolve to a URI relative to a processing environment. For instance the URI reference `#bar` resolves to `http://foo.org/foo#bar` if the XML base in scope is `http://foo.org/foo`. Often the notation `foo:bar` is used, where `foo` is a *namespace*⁷ declaration for URI reference `http://foo.org/foo`, and `bar` is a local name. The combination of both resolves

⁶<http://www.w3.org/RDF>

⁷<http://www.w3.org/TR/REC-xml-names/>

to a valid URI reference `http://foo.org/foo#bar`. The namespace mechanism is important for partitioning vocabularies into modules.

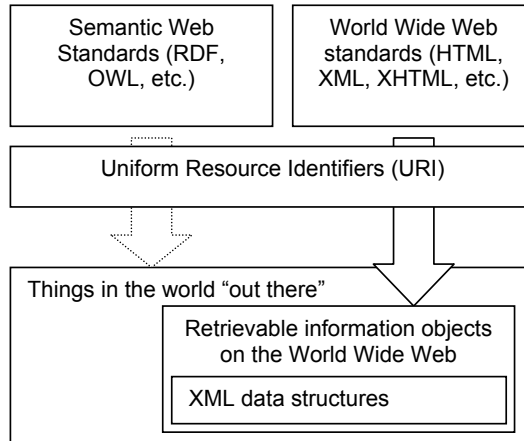


Figure 3.1. Semantic web vs. World Wide Web: one refers to an accessible world consisting of XML data structures and other information objects that can be retrieved, and the other one to an inaccessible one consisting of things “out there” that can only be referred to by way of a symbol.

An RDF description of a resource consists of a set of statements. An RDF statement is a triple with the following components:

subject: the thing the statement describes;

predicate: a property of the thing;

object: the thing the statement says is the value of the property, for the thing the statement describes.

A set of triples that share one or more URI form a labeled graph. The subject and the property value are always URIs that identify something, in the case of metadata a *resource*. The object is either a URI or a (optionally datatyped) literal. A literal identifies itself. Here I will use the notation `(predicate subject object)` for triples (or `(predicate subject "literal")` if the object is a literal). See the RDF specifications⁸ for details on RDF literals and datatyping of literals.

The earlier example resolves to two triples:

1. `(foo:bar rdf:type xmlNameElement)`
2. `(foo:bar rdf:value "Alexander")`.

RDF statements may be reified and identified by a URI. The reification `r` of `(foo:bar rdf:type xmlNameElement)` is:

```
{(r rdf:subject foo:bar), (r rdf:predicate rdf:type),
 (r rdf:object xmlNameElement), (r rdf:type rdf:Statement)}
```

⁸See <http://www.w3.org/2001/sw/RDFCore/>

It is however not prescribed by the standard that the following holds in the other direction:

$$\begin{aligned} &\{(r \text{ rdf:subject } \text{foo:bar}), (r \text{ rdf:predicate } \text{rdf:type}), \\ &(r \text{ rdf:object } \text{xmlNameElement}), (r \text{ rdf:type } \text{rdf:Statement})\} \\ &\models (\text{foo:bar } \text{rdf:type } \text{xmlNameElement}) \end{aligned}$$

Reification can be used to for instance indicate the author of a statement, or to add epistemic qualifications about the statement.

An RDF aware application can correctly parse RDF and manipulates the resulting data structure in accordance with RDF semantics. Of specific interest to us is the combination of RDF and the OWL DL subset of the *Web Ontology Language*⁹ (confusingly, *OWL*) that can be interpreted as a description logic. OWL DL is stored in the form of RDF data structures and processed and manipulated in accordance with RDF semantics.

OWL DL is discussed in the next section. An OWL DL reasoner (e.g. Pellet, Racer, Fact, Fact++, Kaon) is an RDF aware application that applies OWL DL semantics to RDF data structures. Wellknown editors for OWL DL are for instance Topbraid Composer, Protege, and Altova Semanticworks.

There are many features that are common in knowledge representation languages that are missing in OWL. This has led to multiple initiatives to create an additional knowledge representation language for the Semantic Web (e.g. RIF, SWRL) to be used as an alternative for, or in conjunction with, OWL. So far these either fall short of expectations or sacrifice design requirements of the Semantic Web, such as decidability and monotonic logical entailment *with* the so-called open world assumption explained in section 3.4.1.

The need for an alternative knowledge representation language is felt because knowledge engineers are aware of the existence of kinds of knowledge that cannot be properly captured with OWL DL. The introduction of this chapter qualified this knowledge as *ampliative*, and generally defeasible, in character. It is not plausible, given the existing variety in epistemological approaches to this problem, that the World Wide Web Consortium will be able to specify a logic-based language that pleases everyone.

Of special interest for the purposes of this book is introspective, or autoepistemic, reasoning, to be discussed in section 3.4.2. As will be shown in that section, introspective reasoning is already implemented using existing Semantic Web technologies, but as a method of interacting with OWL DL knowledge components, and not as a complete logical language in its own right.

3.3. Knowledge Components and Revision

The idea of reasoning as autonomous process modeled by logical entailment meshes well with the notion of a knowledge base as a black box *knowledge source*¹⁰ interacted

⁹<http://www.w3.org/TR/owl-features/>

¹⁰This use of knowledge source conflicts with the idea of a knowledge source as the origin of the thing to be represented (i.e. a written source or an expert).

with by way of ask and tell operations (specified in detail by Levesque in [187]) using a controlled vocabulary taken from the ontology the knowledge source uses. The black box knowledge source does not itself initiate ask and tell operations. A tell operation expands the belief base with a new proposition. The OWL DL knowledge component as conceived in this book is the static knowledge base, a *belief base*, that can be loaded into a generic OWL DL reasoner, becoming a knowledge source in Levesque's terms. The black box interpretation of knowledge components in Levesque's terms will be used for at least:

1. ontology and
2. abstract "scientific" theories.

Other knowledge sources are however less well-behaved: the represented belief base is defeasible for a variety of reasons, or they resolve a query by performing ask operations of other knowledge sources whose behaviour cannot be guaranteed (for instance a human user). This makes them less freely reusable. At least the following modes of reasoning introduced in chapter 2 belong to this category:

1. abstraction, refinement, aggregation;
2. choice between a menu of actions; and
3. abductive explanation of situations and outcomes of actions.

We can conceive of these other knowledge sources in terms of their interaction with standard OWL DL knowledge sources, and explain defeasibility in terms of *belief revision* on the OWL DL knowledge source.

In linguistics and psychology it is common to think of a *proposition* as the content of a speech act, as something carried by a medium, as the content of a belief, or as something towards which we have a *propositional attitude*. The sources of law are carriers of information, and (therefore) propositions in the linguistic sense. If the relevant knowledge components carry a set of sentences S , a (rational) system applying these sentences as rules is *committed* to the set of sentences $Cn(S)$ where Cn is some operation of logical consequence.

In logic, and usually in computer science, the sentences in S and $Cn(S)$ are not explicitly distinguished: both are considered propositions of equal standing even though the ones not in S have possibly never been believed in the true sense.

In belief revision terms, the set S denotes a *belief base* K , while $Cn(S)$ is the corresponding *belief set* (cf. [134]). The beliefs in K are *occurrent beliefs*, while the other ones are *dispositional beliefs*, i.e. only held when asked about them. $K \models \phi$ if ϕ is in the belief set.

As far as ontology and abstract theory is concerned there is no appreciable difference between occurrent and dispositional beliefs. Ontologies and abstract theories can be freely composed into knowledge sources if designed properly. Given two belief bases K_1 and K_2 , $K_1 \cup K_2 \models K_1$ and $K_1 \cup K_2 \models K_2$. Note that if ontology K_1 imports K_2 , then $K_1 \models K_2$. An ask operation, or query ϕ to the KB is $K \models \phi$. A tell operation ϕ to K replaces K by $K \cup \phi$ if $K \cup \phi$ is consistent.

Definition 1. A *knowledge component* is a set of sentences S that denotes a consistent belief base K of a knowledge source. A *knowledge source* is committed to a consistent

belief base K ; Interaction with the knowledge source takes place through ask and tell operations on K .

For defeasible rules the situation is different. As concluded in chapter 2 defeasible rules generally bridge ontological strata; A set of such rules is the interface to an ontological stratum from another stratum (or other strata). The conclusion of a defeasible rule can be held as a belief as long as it is consistent to do so. The application of a rule as a *if ψ then ϕ* construct through interaction with the tell and ask interface of a classical knowledge source can be conceived of as an interaction that first asks of the knowledge source whether ψ is true and if it is tells the defeasible conclusion ϕ .

A tell operation may *revise* the KB if the KB holds defeasible beliefs and the KBS prefers the told information ϕ over (some) of the beliefs in its own belief base. If $K \models \neg\phi$, K must first be modified by calculating the set of prime implicants of $\neg\phi$, and then retracting at least one belief from *each* prime implicant (cf. [134]):

Definition 2. Let $K \models \phi$ where ϕ is a sentence and K a belief base. A fragment $K' \subset K$ is a *prime implicant* of ϕ in K if $K' \models \phi$ and $K'' \not\models \phi$ for every $K'' \subset K'$. The pair (K', ϕ) could be an *argument* in favour of ϕ , relative to K .

See for instance [215] for methods to enumerate prime implicants.

This procedure is the basis for *belief base revision*, and not the more general belief-set-based belief revision approach (cf. generally [117]) that dominates the literature. The decision whether and which belief to retract is based on an application specific *incision function* (cf. [134]) which implements an epistemic entrenchment theory based on meta-properties of the beliefs. If a prime implicant for instance contains only terminological axioms the KBS is committed to believe anything entailed by them, and cannot revise the knowledge base. If a member of the prime implicant is for instance a *default*, the KBS *by default* believes the entailed conclusions, and may retract the default.

Do note that I advocate making few commitments to the meta-properties of beliefs and the incision function, unless the source of law makes them. As suggested throughout chapter 2 we can distinguish between terminological axioms and axioms from abstract theories on the one hand, and defeasible rules on the other hand. This is a clear commitment. For defeasible rules it makes sense to impose a preference relation on them, if such preferences are expressed in the sources.

The belief base revision approach can provide a semantic basis for more liberal models of user interaction than the classical ask and tell approach allows, but this subject is beyond the scope of this book.

In this book, reasoning with, and importing, and consistency checking of knowledge components is assumed to be implemented in Semantic Web standards-based technology. Semantic Web standards have been designed with this goal in mind, and, in as far as they are successful as a standard, other non-legal knowledge components, that play a supporting role in legal reasoning, will also be available in this form. Section 3.4 discusses the OWL description logic, and the idea behind it.

Section 3.4.2 discusses an approach to representing knowledge that is subject to revision, and interpreting its meaning as a form of autoepistemic, or introspective,

reasoning. This interpretation is based in the observation that there is something peculiar about the way in which the *ask* operation is often implemented in Semantic Web applications. It is a characterization and justification of an already existing approach to using Semantic Web technology in a KBS rather than a proposal of one. In this book defeasible rules are only interpreted in relation to the (non-defeasible) knowledge component to which they belong:

Definition 3. A *defeasible extension to a knowledge component* is a set of defeasible rules, interpreted as recommended or sanctioned but defeasible ask/tell interactions with a knowledge source committed to the belief base represented by that knowledge component.

3.4. Description Logic, OWL DL, and Autoepistemic Interaction

Let us start with characterizing some identifying features of *description logic*, the language used for describing legal knowledge in as far as that is possible. Of interest here is the specific description logic OWL DL.

A DL *knowledge base* consists of a *tbox* – a set of terminological axioms – and a *abox* – a set of grounded axioms about a particular world of named individuals called *assertions*.

The tbox offers a framework for subsumption and classification, and the abox contains statements about particular things or *individuals* in the world. These *things* are denoted by constants. The distinction was introduced by Levesque in [187], in an excellent overview of the functions of knowledge bases, but has gradually changed in meaning somewhat.

The DL knowledge base is intended to be used as a knowledge source as specified in the previous section. It is however common to restrict the possibility of *telling* propositions to the knowledge base to assertions in the abox in normal interaction scenarios.

The description language is composed of:

1. constants denoting terms;
2. constants denoting properties;
3. constants denoting individuals;
4. assertions applying terms to individuals and properties to pairs of individuals;
and
5. terminological axioms constraining the application of terms to individuals and properties to pairs of individuals.

Each term denotes a set of individuals $\{i_1, \dots, i_k\}$, each property a set of pairs $\{(i_1, i_2), \dots, (i_j, i_k)\}$ of individuals.

In this book I use a typographic distinction between **CONCEPTS**, **properties**, and generic constants or variables *a, b* if ambiguity could arise. Alternatively **concepts**, **properties**, **a**, **b** may be used. If I need namespace (ns) qualification, the notation **ns:c** is used. The following is a typographical example of a terminological axiom:

$$\text{FUNCTION} \equiv \exists \text{playedBy}.\text{PHYSICALOBJECT} \sqcap \neg \text{AGENT}$$

Something is a function if and only if it is played by some physical object, and not an agent. The following is an assertion:

FUNCTION(example:c)

The individual denoted by c in namespace ns is a function. Description logics vary mostly by expressiveness. Differences can therefore often be characterized in terms of syntax only. There are however some notable differences between OWL DL and most other description logics that can be explained in terms of a few assumptions underlying the logic, which will be addressed later.

Some readers will be familiar with the modal logic-based semantics of most description logics. To facilitate the understanding of DL axioms for those that are not, I present a translation to FOL based on [35]. See [11, 13, 12] for an overview of efficient tableau methods for description logics, including OWL DL.

OWL DL is an expressive description logic recommended for use as an XML standard by the World Wide Web Consortium¹¹. A new proposal, OWL 2¹², extends the expressiveness even further, but this extension will not be systematically taken into account in this book as it was very new when this book was completed. The description logic embodied by OWL DL can be properly understood as a labeled modal logic, to be characterized by modal frames (cf. [157, 158]).

The set of description logic classes $\mathcal{DL}\text{-CLASS}$ covered by OWL DL is defined by the following rule, where C, C_1, C_2 are classes, P, P_1, P_2 are properties, i_1, i_k are abox constants, and N is a name:

$$\begin{aligned} \mathcal{DL}\text{-Class} ::= & N | \top | \perp | \neg C | C_1 \sqcap C_2 | C_1 \sqcup C_2 | \forall P.C | \exists P.C | \{i_1, \dots, i_k\} \\ & | n \leq P | n \geq P | n \leq P.C | n \geq P.C \end{aligned}$$

The DL class corresponds closely with the notion of *concept* in section 2.3. The construct $\{i_1, \dots, i_k\}$ is a *nominal*, a concept that denotes the set of individuals denoted by the constants in the given set¹³ (see [253] for a detailed discussion of nominals).

Cardinality restrictions (the last four) will not be used in this book. See table 3.1 for an explanation of the syntactic subset used.

In addition, there is a set of properties $\mathcal{DL}\text{-PROP}$ (see table 3.2) used in property axioms that allow us to specify the nature of the intended relation, by restricting domain and range, and specifying i.a. functional, symmetric, and transitive frames. These are not completely FOL expressible. They are also of less interest because property axioms are not used in this book. There are important syntactical restrictions on which frame properties can be *combined* for the same property, and how complex properties may be used in class axioms.

There are class axioms and property axioms. The set of tbox axioms $\mathcal{DL}\text{-TBOX}$ for \mathcal{DL} (see table 3.3) is defined by the following rule:

$$\mathcal{DL}\text{-TBOX} ::= C_1 \sqsubseteq C_2 | C_1 \equiv C_2 | P_1 \sqsubseteq P_2 | P_1 \equiv P_2$$

¹¹Described in <http://www.w3.org/TR/owl-features/>

¹²<http://www.w3.org/TR/owl2-primer/>

¹³I.e. every instance of this concept can be denoted by one of the members of the set.

where C_1, C_2 are classes in $\mathcal{DL}\text{-CLASS}$, and P_1, P_2 any property.

The set of abox axioms, or *assertions*, $\mathcal{DL}\text{-ABOX}$ for \mathcal{DL} (see table 3.4) is defined by the following rule:

$$\mathcal{DL}\text{-ABOX} ::= C(a) | P(a, b) | a = b | a \neq b$$

where C is a class in $\mathcal{DL}\text{-CLASS}$, a, b are abox constants, and P a property in $\mathcal{DL}\text{-PROP}$.

The notion of a DL model is defined as usual on the basis of multimodal Kripke models. $M_{\mathcal{DL}} \models \varphi$ means that a formula φ is true in DL model $M_{\mathcal{DL}}$. The truth definition for this language, and semantic notions such frame, satisfiability and validity are standard (see [24]), subsuming propositional logic (the \sqcup and \sqcap operators are only syntactically different from \vee and \wedge), extended with accessibility of individuals which function as the *worlds* of standard modal logic definitions:

$$M_{\mathcal{DL}}, i \models \exists P.\phi \text{ iff } \exists i' : P(i, i') \text{ and } M_{\mathcal{DL}}, i' \models \phi$$

The notation $\forall P.\phi$ is an alternative for $\neg \exists P.\neg \phi$. Term C applies to individual i is written $C(i)$ as an assertion, which is true if $M_{\mathcal{DL}}, i \models C$.

OWL syntax is based on RDF triples. $C(x)$ for instance translates to the triple (`x` `rdf:type` `C`), and $C_1 \sqsubseteq C_2$ to (`C1` `rdfs:subClassOf` `C2`). The standard description logic syntax used in this book can be automatically generated from OWL using the OWL Latex Tool from <http://owltools.ontoware.org/>. The mapping of description logic syntax to an OWL graph is not always immediately evident, since certain evident notational conventions used in DL notation such as operator precedence and the use of parentheses don't work in OWL's triple-based syntax.

Three important assumptions must be known to understand OWL DL reasoning.

Firstly, the sets of constants denoting terms, properties and individuals are assumed to be disjoint in OWL DL: let us remember this as the *disjointness of universals and particulars assumption*. Terms have instances, individuals do not. The disjointness assumption is a typical condition for the realization of epistemic competence: without it the logic is not decidable. Strictly speaking, this assumption even prohibits us from for instance annotating terms with a human-readable label. The disjointness of universals and particulars assumption prohibits reification of propositions or terms as individuals.

There is good reason to occasionally violate against it, and there are simple solutions for carving up a non-DL problem into DL subproblems (cf. generally [126]). The disjointness of universals and particulars assumption is normal for the description logic world, but RDF does not have it. One doesn't usually combine OWL DL with reified RDF triples. For epistemic-level annotation purposes (recommended for defeasible reasoning, as pointed out in section 3.3) one would however want to do so. Combining OWL DL semantics with annotation of reified triples is in principle not a problem. Note that the proposal to extend OWL to OWL 2 introduces both "entity annotations" and "axiom annotations"¹⁴.

¹⁴See <http://www.w3.org/TR/owl2-primer/>.

Secondly, contrary to many other description logics, the presence of two different constants in the abox does not denote the presence of two different things (i.e. the *unique names assumption* does *not* hold): in OWL we must state specifically that $a = b$ or $a \neq b$. A constant denotes one unique thing, but a thing may be denoted by multiple constants. This means that a set of constants $\{i_1, \dots, i_k\}$ denoting individuals does not in itself give an indication about the set of individuals it denotes: we have to state explicitly that all constants denote different individuals.

Thirdly, OWL DL is based on the *open world assumption*. As will be explained in the next section, where the difference between open and closed world will be addressed, the set of assertions in the abox is often intuitively interpreted as a description of a specific Kripke model of the tbox. Because of the (no) unique names assumption and the open world assumption, this is not generally the case in OWL DL. If K_{DL} is the belief base denoted by a set of OWL DL sentences, $K_{DL} \models \neg\phi$ and $K_{DL} \not\models \phi$ are not generally equivalent, as would be the case under the closed world assumption.

$\mathcal{DL}\text{-Class-SYNTAX}$	$\mathcal{FOL}\text{-SYNTAX}$
\top	$C(x) \vee \neg C(x)$
\perp	$C(x) \wedge \neg C(x)$
C	$C(x)$
$\neg C$	$\neg C(x)$
$C_1 \sqcap C_2$	$C_1(x) \wedge C_2(x)$
$C_1 \sqcup C_2$	$C_1(x) \vee C_2(x)$
$\forall P.C$	$\forall y(P(x, y) \rightarrow C(y))$
$\exists P.C$	$\exists y(P(x, y) \wedge C(y))$

Table 3.1. Correspondence between \mathcal{FOL} and $\mathcal{DL}\text{-Class}$ syntax

$\mathcal{DL}\text{-Property-SYNTAX}$	$\mathcal{FOL}\text{-SYNTAX}$
P	$P(x, y)$
$\neg P$	$\neg P(x, y)$
$P_1 \sqcap P_2$	$P_1(x, y) \wedge P_2(x, y)$
$P_1 \sqcup P_2$	$P_1(x, y) \vee P_2(x, y)$
$P_1 \circ P_2$	$\exists z(P_1(x, z) \wedge P_2(z, y))$
P^-	$P^-(y, x)$
$P \upharpoonright C$	$P(x, y) \wedge C(x)$
$P \downharpoonright C$	$P(x, y) \wedge C(y)$
P^+ and P^*	non first-order expressible

Table 3.2. Correspondence between \mathcal{FOL} and \mathcal{DL} property syntax

Although OWL is quite expressive for a description logic, it does have clear limitations that many other logic-based languages do not have. An OWL DL terminological axiom can only enforce a treelike structure of individuals, since there is no provision for variable binding or for local reflexivity (the SELF concept; cf. for instance [180]).

This obviously also means it is only possible to specify axioms over pairs of constants that are in the extension of some property. A sentence such as $\forall x \exists y C(x) \rightarrow D(y)$ specifying constraints over things that may be unconnected to each other is

\mathcal{DL} -TBOX-SYNTAX \mathcal{FOL} -SYNTAX

$C_1 \sqsubseteq C_2$	$\forall x C_1(x) \rightarrow C_2(x)$
$C_1 \equiv C_2$	$\forall x C_1(x) \leftrightarrow C_2(x)$
$P_1 \sqsubseteq P_2$	$\forall xy P_1(x, y) \rightarrow P_2(x, y)$
$P_1 \equiv P_2$	$\forall xy P_1(x, y) \leftrightarrow P_2(x, y)$

Table 3.3. Correspondence between \mathcal{FOL} and \mathcal{DL} tbox axiom syntax \mathcal{DL} -ABOX-SYNTAX \mathcal{FOL} -SYNTAX

$C(a)$	$C(a)$
$P(a, b)$	$P(a, b)$

Table 3.4. Correspondence between \mathcal{FOL} and \mathcal{DL} abox axiom (assertion) syntax

semantically suspect, and undesirable. This restriction on expressiveness is not one we should try to solve: a *case* to be solved should be characterized by a completely connected graph in the abox, and unconnected entities cannot “influence” each other. If one nevertheless wants to use such axioms, one relates all constants to each other with a generic **related** property: $C \sqsubseteq \exists \text{related}.D$.

The most obvious and most often proposed extensions to expressiveness are directed towards:

modeling more general relational structures: this is usually conceptualized as a freer use of variable binding. This comes down to being allowed to say that two nodes in the tree are the same thing in our domain, i.e. expressions along the lines of $\forall x C(x) \rightarrow \exists y (\text{self}(x, y) \wedge (x = y))$.

introducing ternary or n-ary properties: this is intended to represent tuples with more than three places. This can also be achieved by way of reification of the property as a class, for instance the 4-place tuple **give**(AGENT, RECIPIENT, THEME) to a class axiom $\text{GIVING} \sqsubseteq (\exists \text{agent}.\top \sqcap \exists \text{recipient}.\top \sqcap \exists \text{theme}.\top)$. This however leads to additional requirements on expressiveness of constraints on relational structures, the previous item on this list, since certain features that can be expressed succinctly about properties – for instance symmetry – cannot be expressed in an equivalent form on properties reified as a class.

More general relational structures can be represented with various existing rule languages that can be used with OWL DL. The at the moment of writing new OWL 2 proposal also radically expands the relational structures that can be represented in axioms through the use of role inclusion and local reflexivity (**SELF**) (cf. for instance [180]), provided that the abox is preprocessed to relate all known constants with a generic top level property such as **related**, and adds axioms of the form $C \equiv \exists \text{cSELF}$.

It is generally accepted that n-ary properties are not strictly required for expressive purposes. Reifying n-ary properties, in accordance with the W3C Working Group Note of 12 April 2006 on *Defining N-ary Relations on the Semantic Web*¹⁵, is a preferred solution.

¹⁵<http://www.w3.org/TR/swbp-n-aryRelations/>

Another apparently desirable extension is an extension on the epistemic level, which approaches reasoning on OWL DL knowledge bases with a different set of assumptions than the one outlined above. That the need for this apparently exists is clear if one looks at the methods used for interacting with OWL DL knowledge bases. OWL DL conformant description classifiers usually consist of a description logic theorem prover and a – usually more expressive or in any case different – *query language* like SPARQL (cf. generally [159]).

Abstractly a query language based on OWL DL's satisfiability test can determine whether some (closed) proposition φ is or is not entailed by some DL belief base K_{DL} ($K_{DL} \models \varphi$ and $K_{DL} \not\models \varphi$; cf. generally [159, 157]). More interesting for software development is the derived possibility of determining for which constants i, i' in the DL abox some open proposition $C(i)$ or $P(i, i')$ is true ($K_{DL} \models C(i)$ and $K_{DL} \models P(i, i')$). These open queries, for instance SPARQL SELECT queries, return a result set. These i, i' in the result set are the *known constants* in the abox denoting individuals, to be interpreted with the open world assumption and no unique name assumption in mind, not a complete set of distinct individuals.

If one however *misinterprets* the meaning of the result set in certain ways, one may obtain interesting new results. Facilities like SPARQL SELECT can for instance be used to interface a Datalog-style reasoner to an OWL DL knowledge source in the background. A common *hybrid* design in expert systems is to use a terminological component, an OWL DL knowledge source, in the background, and a datalog-based production rule component in the foreground that queries and updates the terminological component. Other such hybrid architectures replace Datalog with a default logic reasoner or a constraint checker querying and updating OWL DL knowledge sources.

The interpretation these hybrid architectures give to the abox of the OWL DL knowledge source is usually autoepistemic or, with another word, introspective. This observation (taken from [205]) – and a minimal set of constraints following from it – is the basis for the interpretation of defeasible rules in this book.

3.4.1. Autoepistemic Querying and Datalog

While the knowledge a reasoner has about the world is necessarily incomplete, it is reasonable to assume that the reasoner always has complete and accurate knowledge about what it *knows* (cf. for instance the *assumption of closure* in [187]).

When a logic acknowledges that the content of our knowledge base is only a subset of the relevant information that *could* be available about the things in the world that have our interest, we speak of *open world reasoning*. Knowledge representation languages are often characterized by open world reasoning. Since this conception of reasoning is also more suitable to the open ended Semantic Web, OWL DL is also limited to open world reasoning.

The basic idea of traditional logics, including OWL DL, is that the domain being modeled can be abstractly conceived of as a *model*, a set of objects and relationships between them. There are many possible models of the domain: a logical theory, in the case of OWL DL usually a tbox and an abox, admits reasonable models and rules out unreasonable ones. The OWL DL belief base is not a complete representation of the

domain being modeled.

Information and models are separated in this paradigm. The meaning of information is in the mapping between the information and the set of models that are consistent with it. The information tells us that *the* model is in this set. Entailment in this paradigm comes down to checking if a proposed piece of information holds in *all* models that are consistent with the available information. Adding information further restricts the set of possible models.

Many KBS however use another paradigm: that of the database. The relational model underlying databases also models the domain in terms of objects and relationships between them, but makes some radically simplifying assumptions: the only relevant objects and relationships that exist in the domain are those captured by the database information. This is the *closed world assumption*. Secondly, the names in the database uniquely identify objects in the domain, i.e. no two constants refer to the same object. This is the *unique name assumption*.

The result of these assumptions is that there is just *one* model of the domain, the one captured by the available information. There is no distinction between the model and the information about it. This is obviously very convenient, but not always very realistic.

Datalog is a formalization of the database approach in which Horn rules (viz. for instance [220, 205]) are used to capture both the structural constraints on models of the domain and the model itself. The objects and relationships in the model are only those explicitly mentioned in rules.

Rules with empty bodies are often called *ground facts*. A rule with an empty head is often called a *constraint*, and is used to express the fact that a model satisfying the condition described in the body of the rule is not admitted. Usually rules must be safe: a *safe rule* is one where all of the variables in the head of a rule also occur in the body of the rule.

Plain datalog does not include negation, but it is very often extended with a similar construct: *negation as failure* in the body of rules. In keeping with the closed world assumption it is used to draw conclusions about what is not true from what is not in the database, i.e. not known. In this book the notation **fails** is used for negation as failure instead of the classical \neg . Datalog is often used in conjunction with OWL DL exactly because OWL DL cannot draw conclusions from failure.

Restriction of a set of possible models by adding information is not possible in datalog (cf. [220]). Adding non-constraint rules does not restrict the set of possible models, but instead *changes* the model. Constraint rules can be used to make certain *changes* impossible.

An OWL theory describes a *set* of models M , and a datalog program describes a specific model m . If one proposes to interrogate an OWL theory *as if* it were a set of datalog ground facts, which is what often happens in reality if a Datalog-style reasoner uses an OWL DL knowledge source as a quasi-database, one might naively expect to obtain some model $m \in M$, but this is not the case.

Example 2. Take for instance an OWL theory consisting of abox $C(a)$ and tbox $C \sqsubseteq \exists p \top$; In the understanding that x, y are universally quantified variables, a query $C(x)$ binds x to $\{a\}$, and a datalog query $p(x, y)$ produces an empty result set $\{\}$.

This model is however not in M . Or take an OWL theory consisting of abox $P(a)$ and tbox $P \sqsubseteq Q \sqcup R$. Models of this theory would classify the object referred to by a as at least a P, Q or a P, R . An expanded OWL theory consisting of abox $P(a)$ and tbox $P \sqsubseteq Q \sqcup R, Q \sqcup R \sqsubseteq S$ would entail $S(a)$, but a datalog program $S(x) \leftarrow Q(x), S(x) \leftarrow R(x)$ would not, as similar as the Datalog program may seem to $Q \sqcup R \sqsubseteq S$.

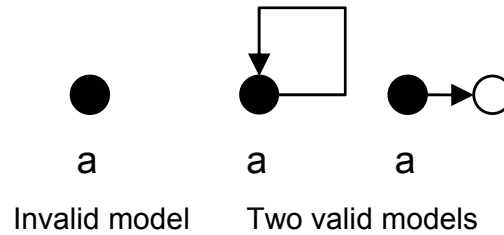


Figure 3.2. Models of C . Solid arrows represent property p . Datalog queries interrogate a model that is not a valid model of C . Instead it represents the settled part of *all* models of C .

The model m interrogated by the Datalog reasoner only consists of the objects and relationships we *know*, i.e. have explicitly represented as a constant in our theory about the domain. It is of course possible to use a set of Datalog rules to update an OWL DL knowledge source with tell operations as if it were a database, but only if the expansions one makes are within the space of possible models. If the Datalog program is carefully handcrafted this may be done successfully, but the maintenance of this Datalog program will be a non-trivial matter.

This mode of interaction between Datalog and OWL DL can be characterized as autoepistemic or introspective reasoning: it gives information about the part of *the* model that is settled, the objects and relationships that are in *all* possible models of the theory. This mode of inference is useful as an ingredient for *ampliative reasoning*, the selective expansion of the part of the model that is settled until we have solved a problem (confer with section 2.4.1 on ampliative reasoning). The expansions of the model we are interested in depend on the problem setting: outside of this problem setting this knowledge is not reusable. Which expansions we prefer largely depends on what is in the settled part of the model, i.e. on the available knowledge.

Note that this autoepistemic characterization only applies to a specific, simple approach to *interfacing* other forms of reasoning to OWL DL knowledge sources *as if* they are databases that return result sets.

This use of Datalog in conjunction with OWL DL is rather crude and inherently unsafe, and may cause inconsistency. More advanced proposals however exist that explicitly take into account defeasibility.

3.4.2. Rules, Defaults, Constraints

As pointed out in [205] there are three familiar uses of closed world reasoning that KBS developers always want to bring in to open world assumption-based knowledge representation languages:

constraints: reasoning about what is missing from the knowledge base to either drive user interaction or to check knowledge availability assumptions; for instance: *Every taxpayer has an SSN number, and John is a taxpayer, but John's SSN number is not known: therefore ask John for his SSN number;*

negation as failure: reasoning about what is *not* known about the domain to draw conclusions about what is not the case; for instance: *there exists no social security record of John, therefore, assuming all social security records are known, John is not a taxpayer;*

defaults: reasoning about what is plausible to assume to be the case, given that our database contains no information in either direction, and when to retract the assumption if the gaps are filled in later. Default rules often have the format¹⁶ *if we know α , and it is consistent to assume β , then conclude γ* , for instance *if John has stated no income to the tax administration, and it is consistent to assume John is unemployed, then assume John is unemployed.*

Constraints and negation as failure were already mentioned in relation to Datalog interfaced with an OWL DL knowledge source as database. Even *default rules* can be used in this hybrid Datalog-OWL setup if the subject of the negation as failure operator can be a negated proposition (cf. [127]). Since OWL DL can represent equivalence and negation, direct interrogation of an OWL DL knowledge base *in effect makes this possible even if the Datalog reasoner has no explicit provision for it.*

KBS using OWL DL knowledge sources often employ some Datalog-based implementation of rules and constraints, including negation as failure, to precisely specify the intended model and to guide information collection, which generally consists of asking for the missing ground facts. In a *carefully handcrafted set* of rules this may work; Multiple such sets, with a control structure guiding their application, may be used as the implementation of a problem solving method. Datalog however has no solution for the existence of multiple, alternative expansions of a model, and the potential interactions between different directions of expansion.

Specialized logics for law often employ some mechanism for *default reasoning*, which does explicitly account for the existence of multiple, alternative expansions, and the potential interactions between default rules. Characterizations of default reasoning are generally based on the notion of *minimal* models. Minimality – when applied to sets of propositions instead of models – was already used as a criterium in section 2.4.1 for *explanation* and in section 3.3 for *argument*. As noted in section 2.4.1, the best explanation of something is often characterized as the disjunction of all *minimal* explanations, ordered by epistemic preference.

Ampliative reasoning is necessary for non-trivial problem solving as section 2.4.1 pointed out. This is however not in itself a justification for any specific approach

¹⁶Following Reiter's default logic (cf. generally [175]).

to representing ampliative knowledge. There is little reason to believe that good explanations are, or should be, generated by these types of rules.

In law we however have to deal with the additional problem that the relevant legal theory, and the legislation to be represented as knowledge, may contain rules explicitly dealing with reasoning on incomplete information. Reasoning with incomplete information is also a key component of dialectical argumentation: without some way to represent competing expanded models of the same settled model there is no way to even conceptualize dialectical settings.

The objective for the purposes of this book is to accommodate some representation of this type of rules in the sources of law that does not permit indiscriminate mixing closed world and open world reasoning.

One can conceive of these rule as plan operators stating advice about what is useful to *add* (without committing to the idea that this is a form of entailment) to the database given a certain objective. What something *is* on the ontological level, cannot – generally speaking – be conditional on how much we know about it: This is a matter of ontological commitment. If a conclusion is based on introspection, it has a different epistemic status than conclusions derived from terminological inference. The epistemic status of the *belief* is a property of the reified RDF statement, and of no interest in this book.

Common extensions of DLs with the function of expressing constraints and defaults can be modeled modal autoepistemic logics (cf. [205]). Autoepistemic logics model for *introspection*, the ability to reason about one’s own beliefs. In these proposals, DLs are extended with an autoepistemic knowledge operator K for known and A for assumption.

The result set returned for previously discussed queries testing $K_{DL} \models C(i)$ or $K_{DL} \models P(i, i')$ for all i, i' used in the abox *is* a correct set of *known* individuals with the requested features, although the sets may turn out to contain duplicates; These queries potentially form the basis for a correct implementation of autoepistemic KC and $K\exists p.\top$ concepts (cf. [72]), and can be explained as a kind of poor man’s autoepistemic logic.

[127] point out that negation as failure *fails* $C(i)$ applied to abox constant i is correctly implemented by establishing $K_{DL} \not\models C(i)$ through a (failed) query. Concept *fails* C has as its extension the complement of KC in the set of constants used in the abox. Moreover, there is no distinction between the K and A operator when applied to a set of abox constants used in a (non-epistemic) OWL DL belief base (cf. [72, 127]): concept AC maps to concept \neg *fails* C , which means that any individual assumed to be in C belongs to the complement of the extension of *fails* C .

As pointed out in [127], Reiter’s default rules *if we know α , and it is consistent to assume β , then conclude γ* translate to $K\alpha \sqcap \neg A\neg\beta \sqsubseteq K\gamma$ from an autoepistemic point of view. The complicated $\neg A\neg\beta$ concept for “it is consistent to assume” evaluates to *fails* $\neg\beta$.

For the representation of the conditions of standard datalog rules and constraints, negation as failure conditions, and “it is consistent to assume” conditions in default rules we can tentatively operationalize three operators *known*, *fails*, and *free* in terms of queries on an OWL DL knowledge source committed to some belief base.

Definition 4. For a constant i used in the abox of an OWL DL belief base K_{DL} :

1. $(\text{known}.C)$ is true if and only if a query $K_{DL} \models C(i)$ evaluates to true;
2. $(\text{fails}.C)$ if and only if not $(\text{known}.C)$; and
3. $(\text{free}.C)$ if and only if $(\text{fails}.\neg C)$.

Interpreting these operators as second-order properties, we may define in OWL syntactic structures outside the DL fragment for the rule in general, and the constraint and default specifically, as meta-concepts **RULE**, **DEFAULT**, **CONSTRAINT**. A simple example is the prototypical default rule:

$$(\text{DEFAULT } r (\text{known}.C_1)(\text{free}.C_2)(\text{assume}.C_2))$$

where **DEFAULT** is the rule class, *assume* is the action of telling the OWL DL knowledge source, and the rule is implicitly universally quantified over the individuals in the abox of that OWL DL knowledge source. Since we are querying on concrete sets of abox constants and pairs of abox constants, the operators could also be used with explicit universally quantified variables $((\text{known}.C(i))$ and $(\text{known}.P(i, i'))$) without significant additional problems.

The term **DEFAULT**, of which r – which violates the disjointness of universals and particulars assumption – is an instance, is defined as followed:

$$\begin{aligned} \text{DEFAULT} \equiv & \text{RULE} \sqcap \exists \text{free.OWL:CLASS} \sqcap \\ & \exists \text{known.OWL:CLASS} \sqcap \exists \text{assume.OWL:CLASS} \end{aligned}$$

The constraint has no *assume* participant.

The rule, which is itself an object, would normally speaking be part of a larger structure, at least a set, specific to a problem solving task. The rules should be conceived of as database update plan operators, to be composed into structures specific for problem solving tasks. If these sets are purposefully composed for a specific role in an application, their application to an OWL DL knowledge source may be known to be safe – i.e. there is no risk of inconsistency.

Inserting back the result of application of an autoepistemic rule into the knowledge base, with information about its epistemic status, can be considered an intentional operation, based on explicit epistemic decision making. The appropriate tactics for this activity on general sets of rules are beyond the scope of this book. Section 3.1 laid out the technical basics of an OWL DL knowledge source revision strategy, based on the calculation of prime implicants. The reason to mention this *belief base* revision approach explicitly, is that this prime implicant calculation feature is already present and is also used for debugging OWL DL inconsistencies.

The implementations and ideas in [205, 72, 134, 168] are compatible with the approach presented here, and extend it in various directions. [175] has implemented Reiter’s default logic on top of OWL DL. If we have to explicitly deal with priority of rules, the approach of [168] is worth mentioning. More advanced solutions allow the assertion of new individuals (existential introduction, or skolemization semantics for existential variables), and more generally allow for a freer use of variables.

Note that there is a great variety of semantics definitions of defeasible logic, some of which can presumably be reduced to selection of an incision function (cf. [134] for the concept of incision function) as pointed out in section 3.1, like the one in [175]. Logics in the same tradition as Reiter’s are the most obvious candidates, and certainly the ones that find application in the field of computer science & law, like [211, 228], are of interest. There are however also a lot of frameworks for defeasible reasoning that are for certain not easily interpreted in terms of database updates and incision functions. See for instance [194] for a systematic inventory of a number of types of semantic characterization of defeasible reasoning, many of which go beyond the idea of theory revision.

3.4.3. Abstract, Hypothetical Objects

Neither OWL DL nor the rules introduced in the previous section allow the introduction of a new object to the *settled model*, as described in the previous section, at least not without accepting that variables occurring in the head of a rule occur in the body only in **free** and **fails** conditions. The set of constants in the abox remains the same. Constraints can only be used to *test* whether the required objects have been introduced to consider a problem solved. To introduce objects we typically have to resort to using OWL DL nominal concepts (e.g. $\text{ABSTRACTCHESSPOSITION} \equiv \{a1, a2, \dots, h8\}$, cf. generally [253] for discussion of the use of nominals in OWL), or we ask the user to supply a new constant with the right features.

One of the reasons for stressing the importance of the distinction between abstract things and the real things they are an abstraction of, is that the extension of abstract concepts is usually clearly delineated (although not necessarily finite). We have no idea how many chess boards and how many chess squares exist in our domain, but on the abstract level every chess position can be explained in terms of the same 1 board and 64 chess squares that we *do know*, and a problem like placing 8 chess queens on a board such that they do not attack each other becomes simply a matter of systematically trying all of them on *the* abstract board. Because the chess KBS is applying an abstract theory that explicitly spells out the relevant abstract objects in the domain, the relevant objects exist in the settled model from the outset. Any *hypothetical object* is necessarily an abstract object.

The automatic introduction of abstract objects through a strategy decided on in advance is important for planning and design systems: in such systems one does not generally ask the users to introduce the candidates, but instead explores the abstract space of abstract candidates implicit in the abstract theories being applied. In many LKBS this is not an issue, since the user supplies the “facts of the case”, and the LKBS limits itself to inferring interesting (terminological and default) propositions about them.

Note that because all hypothetical objects are abstract objects, we can in principle represent the assumptions and closed world reasoning conclusions of any number of agents in the same discourse context, as long as we don’t assert that the things they denote by different constants are the same thing. We can also easily test whether sets of agents have consistent beliefs by trying to equate different abstractions of the same

real world referent. We also don't have to do any bookkeeping outside of the OWL DL model to differentiate belief sets of different agents, since we can add a direct relationship between an agent and the hypothetical objects in his mind.

3.5. About Relevant Knowledge Sources

The notion of sources of law is introduced in 3.5.1, so that the reader will move to chapter 4, which is about law and its function in a social system, knowing what a source of law as intended here is.

These representations of sources of law are used in combination with more general ontologies and theories. In many cases these are perfectly generic: they have nothing to do with law as such. Subjects such as time, wholes and parts, action, do not belong to the field of legal knowledge engineering, but are frequently used in it. Section 3.5.2 reviews some existing ontologies and abstract theories.

We need to distinguish the ontological part of a theory and the defeasible part: most generic theories first require an abstraction step into the ontology, and the theory only yields useful predictions if certain knowledge availability assumptions (usually modeled with autoepistemic constraints) are met.

Section 3.5.3 for instance explains event calculus briefly. Event calculus is both an ontology, a conceptualization of what change is, and an abstract theory for filling in the blanks in an incomplete scenario, based on frame assumptions: if we do not *know* that something changed, event calculus predicts future states based on the assumption that it didn't. This is true for many such generic theories: they consist of an ontological and an autoepistemic component that is more often than not not separated. A theory like event calculus was in the first place developed to explain the use of frame assumptions in planning, not to be a conceptualization of action.

One does not ideally “extend” an abstract theory with legal concepts found in a source of law by importing the abstract theory into the representation of the source of law. The generic theory has a function, it makes a certain epistemic competence possible. The ontological purist distinguishes the following:

1. the concept C_l in the source of law;
2. the abstract concept C_a in the abstract theory; and
3. the autoepistemic abstraction of the concept C_l to C_a , or alternatively, a pragmatic $C \equiv C_l \sqcap C_a$, with the risk that C_l and C_a will turn out to become inconsistent with each other at some point in the future because of changes made to them.

One may however generalize the concepts to which an abstract theory like event calculus can be applied into a *core ontology*. Event calculus is for instance a perfectly viable candidate component of a core ontology, if one forgets about the frame assumptions and their applications. This approach is taken in the LKIF ontology (cf. [153]).

Of importance to chapter 4 are in particular action and intention, and institutions.

3.5.1. Sources of Law

Legal knowledge acquisition only rarely directly appeals to skill, or “epistemic competence in action” as a source of knowledge to be encoded in a knowledge representation. Instead it nearly always directly appeals to written information. Why this is the case is explained in section 4.2, and how the information should be managed and interpreted is explained in section 5.3. Central to legal knowledge representation is the relation between two types of information items: arguments and the legal writings backing them. The knowledge representation represents the meaning of legal writings, and it functions as a resource for generating arguments.

The views expressed in this section on legal knowledge sources are based on [53].

The subsuming category *legal writing* or legal document includes all writings that create a legally recognized fact. All writings required in a court procedure, in legislating, or in exercising a(nother) declarative or proclamative power (granting permits, taking administrative decisions, civil marriage ceremonies, etc.) are potentially legal resources, as well as writings required for compliance with duties to inform (tax application forms, etc.). Very often the writing explicitly includes an argumentation for the creation of the legally recognized fact: more generally it functions as evidence, backing, for its existence.

The legal knowledge source par excellence is the written *source of law*. The source of law is a writing that can be, is, was, or presumably will be used to back an argument concerning the existence of a rule in a certain legal system.

Alternatively, it is a writing used by a competent legislator to communicate a norm to a certain group of addressees. Both the legislator and the user of the source of law understand it as a medium used for communicating the existence of legal norms, as well as auxiliary declarations required for the proper understanding of the legal norms communicated between legislator and user, and *expect* that it can be used to back an argument concerning the existence of a norm or definition in a certain legal system.

Source of law is a familiar concept in law schools, and may be used to refer to both legislators (*fonti delle leggi*, *sources des lois*), legislation and case law (*fonti del diritto*, *sources du droit*), custom, etc. It should be noted that many romance languages make a distinction between the legislator as source of law, by way of speaking or writing, and the written law as source of right(s)¹⁷.

In its broadest sense, the source of law is anything that can be conceived of as the originator of legal norms: for the purposes of this work it strictly refers to communication in writing. There are two main categories of source of law in writing: legislation and case law.

How case law functions as a source of law is addressed in section 6.7.2. Important to note here is that case law as a source of law posits general rules that can be applied to cases, just like legislation. It is not the embodied decision itself that functions as a source of law, but it explicit generalization to a more general rule, i.e. the argument ϕ because ψ is generalized to a rule ϕ if ψ .

The related, but disjoint, notion of a *legislative resource* includes all writings produced by the legislator explaining and justifying legislation. The legislator is a legal

¹⁷Also the Dutch *rechtsbron* is also more accurately translated as source of right.

person: it exists separately from any natural persons and organizations involved in the process of drafting and evaluating legislation. It is the formally correct completion of certain processes, usually dictated by law, that makes the legislator the formal author of a writing, and at the same time identifies the addressees to whom it applies. Obviously, the persons and organizations involved in the process of legislating may produce writings that are clearly precursors or legally required ingredients to the end product. These writings are also included in the notion of a legislative resource, but in this case it is not easy to give straightforward rules for deciding whether they are, or are not to be considered legislative resources. Different jurisdictions will have different theories on this subject.

As a last resort one can look at commentaries, written by independent parties.

Legal knowledge representation aims at an isomorphic representation of the source (cf. [17]), and isomorphism between the results of problem solving and existing written legal argument. The user who consults a system to determine the legality of some situation or action will accept the reference to the source as a backing for an argument if the relation between the argument and source is clarified by reading the source. A finer granularity level is usually better, but if the fragments of text get very small they lose their discourse context. The required granularity of isomorphism depends on the purpose of the system.

As pointed out in [42], deciding on the lexical scope of terms encountered in sources of law is of central importance: in that work an extremely skeptical approach of first representing each separate sentence, and integrating them by way of equivalence statements was advocated. In section 5.3 this approach is worked out in more detail. For maintenance not only isomorphism but also transparent version management (of source and of representation) is essential.

For the epistemic aspect to knowledge representation, legal knowledge engineering can refer to legal theory wherever the sources leave gaps. Legal reasoning is regularly portrayed as especially hard or ill-understood (see for instance [142, 201]), but this is rather caused by the wealth of *written* information – also detailing the many conceptual complications one encounters – available to fill the epistemological gap, and not its scarcity.

Hordes of brilliant minds have shed their light upon the problems of justifying legal decision making over the centuries, for good reasons because the law requires such justification. Few such analyses are available about the reasoning behind medical diagnosis, elevator design, scheduling mail delivery, configuring mainframe computers, etc: here *tacit knowledge* must be uncovered by the knowledge engineer through verbal reporting and observation techniques (see e.g. [257] for an overview of techniques)¹⁸.

Moles in [201] knocks down a straw man, when he criticizes the over-reliance of knowledge engineers on codified law: in at least as many cases commercial LKBS could have benefited from *more* attention for codified law. This criticism also fails to acknowledge *why* LKBS are deployed: we are usually dealing with routine decision

¹⁸This obviously does not mean that there is *no* tacit knowledge in law: the epistemological gap is “closed” when knowledge engineers know enough to implement a KBS. In the case of law it can be almost closed by reading law and academic literature alone. Knowledge engineers obviously also have tacit knowledge, some of which is shared with the people whose tacit knowledge is being uncovered.

processes, and customers are typically not very interested in LKBS that make very original decisions involving especially serendipitous legal reasoning.

3.5.2. Generic Purpose Knowledge

Some domain theories have been worked out in detail in the knowledge engineering field, and can be considered as a given. These qualify as reusable knowledge components, although most of the involved schemas would need to be redesigned to conform to the design principles in this book (with respect to abstraction and separation of autoepistemic reasoning from ontology, or the principles of good ontology design).

Generally speaking, the best worked out domain is *naive physics* (cf. [144]) or qualitative physics. This area includes coherent common sense theories of the time and change (cf. [3, 4]), movement and places (cf. [99]), behaviour of liquids (cf. [143]), etc.

Time and change – in relation to actions and intentions – is quite central to legal reasoning, and merits a separate discussion. This is considerably less the case for, for instance, law and liquids.

Movement and place are relevant, in the field of law known as *spatial planning*: the combination of norms and places has its own logic, but is usually not considered a field of law in its own right¹⁹. The author of this work worked previously on this subject (cf. [40, 45, 46, 41]), in the context of the DURP project on standardization of spatial planning information in XML. They have no central position in legal theory, except in matters of (territorial) jurisdiction, spatial planning, and traffic.

Objects specific to law are however rarely found in physical reality. Most of them exist in the mental domain, where few similar well-developed ontologies exist.

Since the source of law has a commanding position in the legal setting, the nature of documents as expressions of legal knowledge is obviously a core domain, keeping in mind the remark of Makinson in [195] quoted earlier. The *Functional Requirements for Bibliographic Records* (FRBR; cf. [245]) are standard in this domain. In this book, in chapter 5, I follow *MetaLex* practices (cf. [53, 41, 40, 38, 37, 36, 285, 48, 51, 52, 49]), which are i.a. based on FRBR.

Another important area is of course the epistemic domain itself, in chapter 2 (see section 2.2.2) referred to as *folk psychology*. With the domain of intentional things (cf. [248, 148]), things that exist because we intend them to exist, it is a part of the mental domain. Of special relevance is the field of decision theory (cf. generally [137]). Although there is a wealth of scientific literature in these areas, there are no *specific* recommended well-articulated and widely available ontologies of it²⁰.

Norms in law, which clearly also belong to the nonphysical domain, combine several different aspects: they are intentional objects that justify attaching institutional qualifications to things, and they are standards to compare against, i.e. they are evaluative, or express what is better according to the legislator, etc.

¹⁹The field is dominated by social geographers.

²⁰The categories of the mental and the social do lead to a lot of discussion in the development of *top ontologies*, and [202] for instance obviously includes many relevant categories in his legal epistemology.

Betterness appears to be a mental analogy to position: up is better, down is worse (cf. [183]). In law the scale is usually thought of as a binary one, which distinguishes things into legal and illegal. This is an artifact of how the legislator expresses norms: the legislator does not express himself in scales – for instance by grading behaviour like a teacher – but in deontic expressions. The scale is however usually not binary from an *ex ante* point of view: one is considering alternative hypothetical actions to formulate a plan. From an *ex post* assessment point of view it often appears to be binary, and it is often enough to apply it in this way, although we do occasionally need consider the question whether a better alternative was really at hand when a decision was made.

A lot of the academic research in this area is unfortunately under the heading of *preference*: this makes it conceptually a bit hard to consider norms as expressions of *betterness*. Amongst many others the following works have tried to explain norms in terms of preference relations: [138, 268, 262, 43, 44, 89].

3.5.3. Action, Change, and Intentions

Since norms in law can only be explained in terms of intention and decision making (from an *ex ante* point of view) or action and change (from an *ex post* point of view), theories of norms in law cannot be strictly separated from a theory of intention, action, and change. A lot of secondary legal terminology (right, liability, etc.) must be explained in terms of norms and action. In chapter 4 this combined theory will be developed.

Action brings together change and intention. If we consider events to be interesting changes, then actions are whatever it is that relates our intentions and events, honouring Wittgenstein's adagium (in [288]) that the act is in "what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm."

Allen's general theory of action and time (cf. [3, 4]) and the event calculus (cf. [178]) are more or less standard in knowledge engineering, and have often been applied in legal knowledge engineering work. In addition, [151] specifies a standard way of representing time-indexed information in OWL.

The event calculus conceives of events as initiating and terminating fluents, and actions as events. One could say that initiation and termination focus attention to one half of a change. The initiation of one thing is the termination of something else. Event calculus includes frame assumptions that spell out what remains the same unless an event changing it is noted.

The theory focuses our attention on interesting changes against a backdrop of stasis, i.e. it is a typical foreground/background conceptualization, which depend on *granularity*. One could for instance describe the tidal cycle as rising tide (stasis), which turns (event) to falling tide (stasis), which again turns (event) to rising tide (stasis). Alternatively one could describe it as high tide (stasis), which falls (event) to low tide (stasis), and again rises (event) to high tide (stasis). Both are undoubtedly dreadful from a tidal analysis point of view, but a mention of the tide in a legal case will very likely pick one of these two perspectives. What the events are and what stasis is depends on what we are interested in (cf. generally [86]).

The law may obviously choose different levels of granularity, which may be incompatible (i.e. they don't reduce to each other), and one should respect the original formulation. Note that both events and stasis may take an unspecified amount of time, and both may be reduced into events and stasis at a lower level of granularity.

Our conception of what changes when has a large influence on how we reason about action, and about what ought to be done. Many logics of agency treat actions as instantaneous changes of state occurring at discrete time steps (i.e. *Hoare logic*, *dynamic modal logic*). It goes without argument that this is a superficial characterization of the nature of change (compared to [86]).

Its main feature is that this simple conception is computationally very powerful: the behaviour of computer programs is one of the domains for which this conceptualization works satisfactorily²¹. While even this theory is undoubtedly useful for legal knowledge engineering in some cases, in particular on the epistemic level, as a description of reasoning itself, one should be very wary of the conclusions one can reach using this theory about other domains, in particular the lawfulness of actions, because it embeds an important and very questionable assumption, as noted in section 4.7: the logic, instead of the law, may dictate whenever the *deadlines* occur.

The combination of Hoare-style logics with deontic logics of norms is an excellent example of the risks involved in mixing the autoepistemic assumptions of an abstract theory of one domain with those of another. This point will be discussed in more detail in section 4.7.

If we assume that the changes brought about by action are always intended, intention plays only a minor role. Although all actions are intentional, i.e. intend to achieve some change, they may however also fail to do so, or cause unforeseen side effects. Since the law distinguishes between intent, negligence, recklessness, etc. we do as a rule have to take into account the intentions of an agent, and we therefore have to commit to a “folk psychology” model of how the agent makes decisions.

Closely related to action and intention are social and institutional reality: these exist because they are intended to. These have been the subject of extensive philosophical debate (viz. for instance [248, 148, 164, 221]), but are very conspicuously glossed over – perhaps as too arbitrary to axiomatize them – by most ontologies. Institutional reality is one of the main subjects of chapter 4, and will be explained in terms of the *constitutes* and *executes* properties, that function as bridges between three different ontological strata.

3.6. Conclusions

Section 3.3 postulated the possibility of using knowledge components. If such a knowledge component represents the kind of knowledge that can be represented in OWL DL we take a *reasoning as process*, or *design* stance towards it. This is the case for ontologies and for abstract theories.

OWL DL knowledge components can be composed into structures of components through the import mechanism. An OWL DL knowledge component does not interro-

²¹That is, it works as long as everything of interest runs by the same system clock.

gate other OWL DL knowledge components, but directly imports them, if the resulting *belief base* is consistent. Composability is in principle only restricted by the constraint that the result must be consistent. From the point of view of impact analysis however, some import trees are better than others: For the purposes of this book, the representation of a source of law does not import generic purpose knowledge beyond the properties and concepts identified in this book for that purpose.

An OWL DL knowledge source is created by importing the OWL DL knowledge component into a conformant reasoner with a standardized query interface. A peculiarity of common query interfaces, like SPARQL, is that they permit an autoepistemic interpretation of the OWL DL knowledge source.

The interrogation of OWL knowledge bases by a Datalog-like rule language (cf. section 3.4.1) through a standard SPARQL query interface is a common setup for KBS. Because the rule language is not integrated into the semantics of the description logic, the resulting formal system is hard to formalize; Section 3.4.2 argued that this setup performs the essential function of autoepistemic reasoning. Three important such functions are enforcing constraints, inferring negation from failure, and applying default rules. Since there is a large variety of accounts for how and when such rules must be applied, we take a reasoning as decision making stance to it, and do not commit to an approach. The rule can be conceived of as a plan operator, to be used as part of a plan decided on in advance. In principle, it is possible to implement such rules in Datalog, if properly stratified through for instance the use of problem solving methods. It is also possible to use generic approaches to applying defeasible rules; Section 3.4.2 pointed out that Reiter's default logic is for instance implemented as an interface to OWL.

This chapter introduced two types of rules, the `DEFAULT` and the `CONSTRAINT`, whose semantics as a database update to the OWL DL knowledge component is relatively well-understood when considered in isolation, that will be used in this book. A set of these rules is a defeasible extension to a knowledge component. The knowledge component thus consists of:

1. a set of terminological axioms in OWL DL; and
2. a set of default and constraint rules.

A general principle is that it is only through the defeasible extension that one interrogates description logic knowledge components, that only the defeasible extension is used to drive interaction with human users, and that description logic components only import description logic components.

The focus of this work is the representation of sources of law as knowledge components. For the purposes of LKBS maintenance, one ideally maintains a clear mapping between knowledge components and the sources of law they are based on. The source of law is a writing that can be, is, was, or presumably will be used to back an argument concerning the existence of a rule in a certain legal system (cf. section 3.5.1).

In the next chapters some things will be taken for granted. There is no reason to dwell on the decomposability of things into their parts, or what effects this has in terms of inheritance of properties and relations from part to whole or vice versa. If article 2 of act *a* is changed, then *a* is also changed. If the date of publication of *a* is

t , then the date of publication of article 2 of act a is also t , etc. Time and space will also be taken for granted. Section 3.5.2 lists a number of these very general knowledge domains, that are so generic that it serves little purpose for this work to dwell on their complexities if it is of only ephemeral interest to interpretation of sources of law. Only the essentials, as proposed in the next two chapters, matter.

To bridge *ontological* strata the **constitutes** and **executes** property will be regularly used in this book, for reasons explained in chapter 4. The presence of certain properties tells us that a logical formula is not an ontological one. The major purpose of the use of *ontological stratification* in the following chapters is to methodologically align a content-based criterium – what does a the rule say? – with a technical concern – is the rule defeasible?

Chapter 4

Agency and the Law

4.1. Introduction

Society uses a wide variety of sources of law, and uses many labels other than legislation to describe sets of written rules: quasi-legislation; administrative rules; codes of practice; guidance; guidance notes; policy guidance; guidelines; circulars; framework documents; outline schemes, and statements of advice. A commonality of these documents is that they are understood to postulate or describe norms that guide or mandate conduct in a given type of situation. This guidance can pertain both to physical behaviour and to decision making processes.

Before one can turn to the question of how to represent legislation and similar documents, one must consider the role these documents play in guiding behaviour.

This chapter deals with the relation between these formalized rules and behaviour. It does not address the problems legislators encounter in managing a large body of such rules, the solutions they have found for these problems, and the consequences these have for knowledge engineers who try to represent the meaning of these rules in logical form: this is the subject of the chapter 5. In this chapter a number of legal theoretic devices are introduced, and positioned as practical reasoning devices in the context of *planning* and *plan recognition* problems: since this context is missing when we consider sources of law by themselves, a representation of the source of law will often omit this information. Chapter 6 reintroduces some of these notions in the context of an *intended normative order*.

One of the major objectives of this book is to consider the problem of representing legal rules while making minimal epistemic commitments to the way these legal rules are used. At the same time we concluded in chapter 2 that we have to account for two very different uses of legal rules:

Planning to perform a task one set oneself; in the case of law to bring about beneficial legal facts, while avoiding detrimental ones; and

Situation and action recognition to infer one's own (legal) position, and what relevant (legal) facts others bring about, are going to bring about, and unsuccessfully attempted to bring about.

Both generic types of intelligent problem solving behaviour are often influenced by legal rules, but the legal rules are of course not the only source of knowledge taken into account.

Planning involves generating and comparing alternative plans, the execution of which involves behaviour, while situation and action recognition involve generating and comparing consistent alternative explanations of observed behaviour. Both subjects are not of specific interest to the field of law, but the realization of interesting problem solving competence will invariably involve strong commitments, and often assumptions, about the epistemic framework in which legal rules are used. Such commitments to epistemic frameworks are often deeply embedded in legal theory, and in representation techniques based on legal theory.

This chapter surveys a number of important legal theoretical concepts used in legal knowledge engineering, in an attempt to separate epistemic reflection on the use of knowledge of the rules from the reusable rules proper. The law as an *institution* is a

central concept.

A subject that is carefully avoided in most of this chapter is the notion of *reified* legal rules with their properties, and the properties of the documents that contain these rules. Instead this chapter discusses the logical form of constitutive and institutional rules.

4.1.1. The Institutional Perspective

This book takes law to be an institution whose primary purpose is to create normative order by way of formalization – the sources of law specify a *formal, institutionalized normative order* (cf. [192]) – usually wherever spontaneously arising, informal normative order fails to achieve the desired results.

The following rationales are usually associated with the formalization, i.e. the writing down, of rules:

Accumulating Knowledge Accumulated knowledge from conduct in the past is written down in the form of rules to guide conduct in the future. The rules set out the relevant criteria, guide collection of information, decrease the amount of mistakes, and generally allow relatively stupid people to solve complex social problems.

Consistency Writing down rules supposedly encourages consistency, fairness, equality of treatment of persons, groups, and organizations in different places and at different times. rules limit the discretion of the decision maker in treating specific persons, groups, and organizations differently, and therefore reduces bias and corruption. Written rules also make behaviour of others predictable, which reduces conflicts.

Democracy The process of writing down rules allows for greater public involvement than the mere making of a decision. The use of rules is in this sense a precondition for effective democracy on a large scale.

Legitimacy Written rules contribute to the perceived legitimacy of decisions, because of the reasons above, and allow the decision-maker to cite the source to justify the decision.

The rules create normative order, but they are not norms.

Much of the terminology to describe institutions has been strongly influenced by the work of Searle (cf. generally [249, 248]), although the notion of institution and the closely related notion of constitutive rules were introduced earlier. Searle describes the distinction between two types of rule, being *normative* – or regulative – and *constitutive* rules, as follows:

Some rules regulate antecedently existing forms of behaviour. For example, the rules of polite table behaviour regulate eating, but eating exists independently of these rules. Some rules, on the other hand, do not merely regulate an antecedently existing activity called playing chess; they, as it were, create the possibility of or define that activity. The activity of playing chess is constituted by action in accordance with these rules. Chess has no existence apart from these rules. The *institutions* (emphasis mine) of marriage, money, and promising are like the institutions of baseball and chess in that they are systems of such constitutive rules or conventions ([249], p. 131).

This distinction has been taken up by many in computer science & law (viz. [29]), although often with a very limited application: constitutiveness is taken to be a characteristic of a limited type of legal rules or legal acts. In legal theory in general there is a bewildering array of conceptions of constitutiveness (cf. for instance [199]), some of which are irreconcilable, and this book will add yet another one.

Both the regulative and the constitutive rule are to be distinguished from the knowledge representation rules of chapter 3 that merely generalize past observations in the hope that they generalize to future ones. The regulative and the constitutive rule on the contrary only affect events that happen after the rule comes about, and application of the rules by an agent is presumably the force that makes the consequent of the rule happen. The regulative rule guides behaviour, while the constitutive rule guides interpretation of events in the terms of an *institutional reality*.

The regulative function of law is generally taken as the defining one of the legal system in positivist legal theory. This is for instance apparent in Kelsen's notion of a *Grundnorm* (cf. [172]), which is clearly of the regulative type, or in Austin's adagium (cf. [10]) – later developed by Hart (cf. [140]) – of law as the command of the sovereign backed by the *threat of punishment*. Normative rules embody the primary purpose of law.

Section 4.2 of this chapter will kick off with a discussion of constitutive rules and institutional facts, since the regulative function – even if it is in the end the really important one – is secondary to the function of constitutive rules. To describe some uses of constitutive rules we have to spell out the relation between someone's decisions and someone's behaviour. In law, this connection is made through the concepts of intention and action, discussed generally in section 4.3, and applied to law in section 4.4.

Section 4.4.1 introduces a number of systematic relations between the rules of the institution, and the types of knowledge representation rules introduced in section 3.4; This is an application of the idea of coherent ontological strata versus defeasible mappings between strata introduced in chapter 2 to institutions as coherent systems versus an uncontrollable environment.

Law *institutionalizes* normative order by way of institutionalizing recognition and evaluation of conduct. The institution and the normative order it institutionalizes are however two separate things. This point of view will be worked out further in sections 4.2, 4.5, and 4.6. Attention to only one of these two aspects, or the insistence that rules are to be *either* conceptualized from an institutional point of view or a regulative one, will lead to unbalanced knowledge representation.

4.1.2. Norms, Normality, Normative Rules, and Agent Roles

Normative rules, discussed in section 4.6, occur in three main variants in natural language: obligations, permissions, and prohibitions. The use of specific auxiliary verbs is however not at all indicative of whether some proposition is normative or not. Both a legislator and a sociologist can utter, for example, the sentence "Tanks adjacent to the hull are not used to store fuel oil", and in neither case the sentence would be semantically ambiguous. In the first case, the legislator sets forth a rule prescribing

how a tank adjacent to the hull should be used. In the second case, the sociologist verifies a regularity, analyzes the relation between the norm and social reality. A legal normative rule is normative by virtue of its function and the way it came about, not of its propositional content.

Section 4.6 takes the position that normative rules are in principle standard constitutive rules, whose conclusion – allowed or disallowed – however connects to an understanding of the world in terms of better and worse actions.

Clearly, normative rules only address conscious human choices. The eruption of a volcano may very well be undesirable, but it makes no sense to try to tell it ought to behave. It is also not very sensible to prohibit making mistakes. Normative rules are information items intended, by their author, the legislator, to influence other peoples choices. The normative rule describes a choice, and the preference the author holds for one of the involved alternatives, and it imposes that preference on the addressees by way of some social mechanism: it is this social mechanism that results in some scripted behaviours and transactions becoming a *norm*.

This social mechanism is generally speaking of relatively little importance to knowledge engineering. Section 4.5 gives this question more attention than most literature in the field, and also touches on the perspective of the legislator (who will return in section 6.7.3 later).

Rules in the sense discussed, normative or otherwise, address agents by *role*. The agents addressed by the rule are described, and not identified. The rule applies to you because you are (recognized as) a citizen, driver, legislator, civil servant, judge etc.

Not every command is a normative rule. A command that addresses a role filled by one person at the time, for instance the monarch or the prime minister, is still a normative rule. A command that directly addresses *you* cannot possibly be one.

Agent role, discussed in section 4.5.2, is the most useful way in which legislation is organized. Tax law addresses taxpayers, criminal law addresses state and citizen, private law addresses interaction between citizens, labour law addresses interaction between employer and employee, etc. Also the attribution of legal powers or competences, addressed in section 4.7.3, is mediated through the role. Reciprocal recognition of actions by persons depends on the the adoption and attribution of agent roles.

It is the understanding of situations in terms of the rules, which connects a set of preferences and expectations to persons through the agent roles attributed to them.

The deontological interpretation of normative rules in section 4.6 is based in the assumption that people avoid the risk of punishment, i.e. consider options that are allowed better than those that are disallowed. This interpretation can be used to plan your own actions, but also to predict the behaviour of others to your advantage, the theme of section 4.8. The other can also be a collective, for instance everyone, as section 4.8.3 explains.

Not only normative rules create preferences and expectations, however, as section 4.5 explains. All rules will lead to habit formation.

4.1.3. The Problem of Agency

In the world around us, we see that formalized norms are usually adopted, provided they are known. There are several explanations for violation of norms, assuming in the first place that it is based on wrong choices:

Unwillingness The agent prefers the wrong alternative. Conditioning (reward and punishment) is the universal solution for this problem, and according to some influential voices the fundamental function of law.

Ignorance The agent failed to understand the law, failed to conceptualize possible choices, or failed to foresee possible outcomes of a choice. If the agent drives on the wrong side of the road, for instance, without the intent to cause a crash, the problem is apparently that the agent failed to foresee the likely crash that would follow. Law can help solve this problem by commanding people to drive on one side – in this case the side is irrelevant – of the road.

Force Majeure In some cases an agent may be forced to choose contrary to the rules because a legal alternative was not available. This case is usually either not considered a violation, or it is a special case of violation without adverse legal consequences.

Non-compliance with the rules is not always an issue of unwillingness. As section 4.3 points out, describing the behaviour of others presumes that you understand their intentions. Since these intentions are not directly observable, but can only be guessed at based on observable aspects of behaviour and preferences and knowledge that you ascribe to the other, applying the rules to the behaviour of others is a complex reasoning activity.

Many of the common complications arising from this, and the assumptions we must minimally make about agency in order to represent normative rules and apply them in planning and action recognition, are addressed in section 4.7 and section 4.8, which discuss situations and actions, deadlines, states and change, normative positions, powers, jural relationships, and legitimate expectations.

4.2. Institutions and Rules

Institutions are, according to a common definition, “structures and mechanisms of social order and cooperation governing the behaviour of two or more individuals”, for instance law, marriage, money, democracy, the marketplace, church, school, etc.

Law is an institution whose primary purpose is to create normative order by way of formalization – i.e. a *formal, institutionalized normative order* (cf. [192]) – usually wherever spontaneously arising, informal normative order fails to achieve the desired results. A similar characterization is already found in the work of Geiger, who speaks of law as a mechanism to create normative order monopolized by a central authority (cf. [119]). The notion of law as an institution, based on so-called constitutive rules, has been a productive concept in legal theory and computer Science & law, in particular to justify the use of logic programming rules to represent law (cf. generally [29, 199]).

An *institution* is a collective intentional or *social* entity, i.e. an entity that exists merely because a collective (i.e. group of natural persons) recognizes and intends its existence. The institution itself is a good example of *intention* applied to the existence of artifacts.¹

The structures of the institution are defined by the institutional facts that make up the institution, and its mechanisms of change are the constitutive rules that specify what *constitutes*, or counts as, an institutional fact. Conversely, the institutional fact has a constitution base (following Hindriks in [148]), which consists of the application of a constitutive rule to the *constituting facts*, which are *brute* facts, yielding an institutional fact.

Definition 5. A *constitutive rule* of an institution derives an institutional fact from one or more other constituting facts, at least one of which is a brute fact.

Definition 6. A *brute fact*, relative to some institution, is a fact that has no constitution base in that institution.

Brute facts are pre-existing and external to the institutional reality constituted by the rules. This does not however mean that they are in any sense “natural”, or non-institutional. The institutional facts of one institution can be the brute facts of another one. If we for instance state that checkmate constitutes winning, checkmate is an institutional fact of chess, and winning is an institutional fact of games. If we consider “winning the chess game” to be part of the institutional language of chess, the relation between checkmate and winning can no longer be considered truly constitutive, unless one expands constitutive to include any terminological axiom, which I will not do. On this point this book follows observations by Hindriks on Searle’s work (in [148]) and not Searle (cf. [249, 248]), who does insist on separating social reality from a pre-existing one. Hindriks identifies these rules as *essential rules* of the institution: the ontology of the institution. An institutional fact that merely rephrases other institutional facts using the institutional ontology shares the constitution base of the institutional facts constituted by brute facts.

Definition 7. An *institutional rule* of an institution derives an institutional fact from one or more other constituting facts, all of which are institutional facts.

The distinction between institutional and brute facts is very similar to the role of the Breuker’s *legal abstract model* (cf. [63]): a layer of “legal” concepts and relations built on top of a large layer of commonsense knowledge ([266], p.57). In Breuker’s original conception this world knowledge could ground different institutional normative orders in a single consensus reality. This is an attractive proposition for the purposes of comparison of two institutional realities that do not intersect at all, as both would presumably be grounded in the same consensus reality, giving us something concrete to compare.

The (institutional or brute) fact is simply a reified statement about something. One may think of the constitutive rule as a simple institutional qualification of something

¹See my earlier observations about the interchangeability of the planning and design problem type in section 2.4.

– the predicate or concept is institutional – or as giving rise to a new institutional thing. This is usually a question of modeling style and requirements, and not a quality inherent in the constitutive rule.

Example 3. In an auction, raising a hand constitutes a bid; This may be represented as $\text{RAISEHAND} \sqsubseteq \text{BID}$ or $\text{RAISEHAND} \sqsubseteq \exists \text{constitutes.BID}$, or a more detailed representation which takes into account the auction setting, the participation of some person in it, the timing of the raising of the hand, and – consequently – the amount of the bid, etc. We can speak of the institutional concept BID, some institutional object o , which is institutional because it is of type BID, and the institutional fact $\text{BID}(o)$.

From an ontological point of view the act of raising one’s hand and the act of bidding a certain amount cannot be one thing, because they do not share essential properties. The raising of a hand is a physical event, while the bid is an institutional, and therefore social, event.

Proposition 2. The *constitutes* (inverse: *constitutedBy*) property; Every institutional thing is constituted by a non-institutional thing:

$$\text{INSTITUTIONALTHING} \sqsubseteq \exists \text{constitutedBy}.\neg \text{INSTITUTIONALTHING}$$

Note that since this *subjective* restriction applies to one and the same institution, and not institutions in general, this is a design pattern for ontology represented in OWL rather than a terminological axiom.

The *constituting act* of raising a hand to make a bid and the supervenient *institutional act* of making a bid are in the example inextricably linked to each other. The same is the case for a constituting act of taking that constitutes a theft. The case is often less clear-cut when we are talking about states, objects, substances, etc. In many cases we see that although the constitutive rule states that some brute state constitutes some institutional state, the act to initiate or terminate that state is labeled constitutive. This is not entirely surprising, since we are dealing with *intentional* things brought about by actions.

A constitutive act in this sense is an act of representation: its *function* is standing for, or representing, something else. Functions of things are not inherent in the thing: they are assigned by (intentional) agents, and the recognition of the function depends on the observer.

This creates an ambiguity, however. We raise our hand at the auction *because* we intend to make a bid: the action has no other rationale. The thief on the other hand has no intent per se of being labeled a thief: he merely takes that risk. To resolve this ambiguity, we have to distinguish the constituting act from the constitutive act. This is the subject of the next section.

4.2.1. Constitutive Acts, Speech Acts, and Intent

In law, a legal act² is an act that creates a legal fact i.a. *because* it was intended by the actor to do so. In analogy, consider a *constitutive act* an act *intended* to constitute

²*Rechtshandeling* in Dutch doctrine, *Rechtsgeschäft* in German doctrine.

an institutional act. A constitutive rule may specify that an act (for instance raising one's hand) constitutes an institutional act *because* it was intended to do so: the constitutiveness of the act is a condition of the constitutive rule. To unambiguously create the institutional fact we have to make a demonstrable declaration of will, and, conversely, to prove the institutional fact one must provide evidence of the intent to produce an institutional fact.

Definition 8. A *constitutive act* is an act intended to constitute an institutional act.

As stressed by for instance Jaap Hage in [132] *intent* has a central role in bringing about institutional facts, but only in its relation with *action*: the institutional fact is brought about by performing a constitutive act, and – normally speaking – an action presumes intent to bring about the product of the action. To describe behaviour in terms of actions is to presume intentionality to produce the products of the action: this is why actions can fail. If a driver turns right we are talking about an action, and if the driver happens to hit a bicyclist while doing it, this is an event, and it is an (unintended) side effect of the action which is not part of the essence of the action itself. Without the action the event could not have happened, but hitting the bicyclist is not the action itself, which was aimed at turning right.

Raising one's hand at an auction constitutes a bid. Raising one's hand to a friend may however also be an attempt to draw attention, and conventionally constitutes a greeting. Raising one's hand to a friend walking in on an auction in the central lobby of a cruise ship you were watching constitutes a greeting, but may be *recognized* by the auction master as a bid, which is an unintended side effect of the action. The auction master's problem is to determine who participates in the auction, and the setting is in this case not conducive to this determination.

Legal theory has a whole battery of concepts to deal with intent, the failure to bring about what was intended, and the fact that not all effects of actions are intended ones: declaration of intent, intent and conditional intent, attempt, recklessness, negligence, mens rea, etc. Still the operative principle behind constitutive rules and institutional facts is that people to a large extent have control over what institutional facts they bring about: this is the whole point of trying to regulate behaviour by way of declaring sanctions in advance.

A *constitutive act* is an act of representation that creates some institutional thing that wasn't there before, by virtue of the intent to create the institutional product of the act alone.

Collective intentionality – Searle's original requirement for performing constitutive acts – does not necessarily presuppose that collectives *are* intentional: it is sufficient that its members can make a distinction between *I*-intentions and *we*-intentions, since the interpretation of institutional reality happens in each individual mind without recourse to a collective one. I intend to undertake some action because I expect that *we* interpret the action as representing something else. Searle specifically defends the thesis that it is unnecessary to explain how *we*-intentions reduce to *I*-intentions. This is an *internalist* point of view, and also the one we would usually take on behalf of (the user of) an expert system.

Collective acceptance, an alternative concept introduced by Hindriks (cf. [148]), derived from Searle's collective intentionality (cf. [248]), can be taken to mean actual acceptance by at least one person besides the actor. It answers a variant on an old philosophical question: when nobody's around, can your actions be truly constitutive, regardless of your *we*-intentions? According to Hindriks this is only the case if they are recorded and later accepted as constitutive by the required audience that accepts the act.

This distinguishes the constitutive act from other acts of representation inherent in speech acts: if one for instance requests someone to do something, the request represents an attempt to make someone do something, regardless of whether it is *accepted* by anyone. Similarly, a promise to do something represents an intention to do something, and only when it has a constitutive character gives rise to a duty to do that thing. Think for instance of a threat: a duty only follows from the threat if there is an audience for whom a threat has the conventional effect of producing a duty, which is nowadays not normally the case. On the contrary: for most people the threat constitutes a crime. The only thing that distinguishes threats from promises is that the thing you promise to do is not appreciated by the target audience.

Although there are some types of speech act typically associated with producing certain conventional effects – declarations, directives, promises, etc. – constitutive acts should not be equated or confused with specific types of speech act.

Secondly, the fact that some act constitutes another act does not mean that the *constituting* act is also *constitutive*: this is in principle only the case if the actor *himself* intends to perform the institutional act.

The constitutive rule can always be applied by bystanders to someone else's actions. In some cases application of a constitutive rule involves *ascription of intent* to someone else's actions, and therefore also the explicit representation of intent, while in many other cases it does not. The representation of constitutive rules should also include the possibility of representing intentions. The next section proposes a simple representation.

4.3. Intention and Action

Section 2.4 distinguished between the internal and external perspective, and the ex ante and ex post perspective on decision making. The ex ante and internal perspective of the decision maker who considers a *plan*, is very different from the ex post and external perspective of the observer who describes someone else's behaviour in terms of *actions*, i.e. ascribes a plan to others based on their behaviour.

This relation between observable behaviour and plans is mediated by the concept of *intention*. It has two distinct senses, depending on whether we take an internal or an external perspective. That is, we intend to do things and we do things intentionally. The relation between what we intend to do and what we do is too weak to reduce these senses of intention straightforwardly to each other (cf. for instance [58]). This is also recognized in law, which uses certain concepts and fictions to deal with this problem.

Ideally, we want to connect intentional action, and its conceptualization in law, to the planning problem type of section 2.2.4, without however committing to a specific

approach to *automated* planning.

To understand what intention is, it makes sense to focus on its *ex ante* sense: Bratman gives good practical and philosophical reasons for the *methodological priority of ex ante intention over ex post intention* (cf. [58]). *Ex post* ascription of intention can simply be considered an explanation of someone else's behaviour in terms of his (*ex ante*) intentions.

Alternatively, we may consider the *ex post* recognition of intent as the primary one, and simply consider the recognition of your own intentions as a special case of recognition of intentions in general: this interpretation is more in line with Newell's position that knowledge is something that is ascribed to intelligent agents to explain their behaviour (cf. [209] or section 2.2.2).

The distinction is central to the theory of mind (cf. [121]): Bratman's internal, *ex ante* theory is representative of the introspectionist or *simulation theory approach to the mind*, while the Newellian external, *ex post* theory is representative of the *theory-theory approach to the mind*. We are either simply applying an acquired *theory* that we use to understand behaviour in general to ourselves, and have no privileged introspective access to the state of our own mind, or we do have such introspective access, at least partially, and have the ability to imagine ourselves in somebody else's shoes by way of a kind of *simulation*.

For the purposes of this book, the acquired theory approach to intention is the best fit. We conceive of the ontology of the mental domain in terms of metaphors and abstraction, and expect for that reason that the way we conceive of planning is a derivative of the way we interpret behaviour.

Intention is clearly linked with planning and action, from both the simulation theory and acquired theory perspective. Observe that intention cannot be simply explained in terms of preferences, desires, goals, etc. One does not usually say that one has the intention to become a famous football player just because one dreams of being one, and one does not generally intend to have the same cake and eat it even if one desires to do so. Intentions *are* closely and inextricably linked to specific actions: intention is the commitment to a plan one decided to execute or are executing, or to a design one decided to implement or are implementing (cf. generally [27] who have a similar view on intentions).

Cohen and Levesque for instance list the following requirements on a satisfactory representation of intentions in [85]:

1. Contrary to desires and preferences, intentions held by an agent are consistent with each other and with the agent's beliefs.
2. An intention poses a problem for an agent: the agent must have a plan that he believes realizes it under certain conditions, and he must be committed to executing that plan.
3. The agent monitors the success or failure of his attempts to realize his intentions: failure constitutes a new problem.

One has the intention to do something when one is doing it. Whether the intention started just before execution of the plan, when one "committed" to it, or just after it when one "recognizes" it, is for engineering purposes of minor impact.

A description of behaviour in terms of actions is thus grounded in the recognition of behaviour as a *plan execution*. It describes a structure in two different strata: a plan in the mind of the agent that he is attempting to execute, and observable behaviour of the agent that succeeds or fails to execute that plan. It is based on the presumption that the agent, whose reasoning capabilities are limited, *must* commit to a plan and remains committed to it for a longer period of time. The agent only takes decisions once he has completed an attempt to execute a plan, either because the plan failed, or it succeeded, or because something interesting came up.

The acquired theory approach, and law, presumes that we can ascribe plans to others with reasonable accuracy based on their behaviour. Note however that the agent can interleave the execution of multiple plans, and that some plans may involve monitoring things. One may for instance intend to buy something as soon as one gets the chance: plan execution starts immediately, with a monitoring activity, but it will be difficult for other agents to recognize the plan execution. This raises some questions for the acquired theory approach, but not for the relation between plans and actions as long as one is willing to consider passive monitoring actions as part of a plan.

Another case which raises more general questions is when the plan requires us to schedule the start of its execution. If the public transport planner tells me that I have to leave in five minutes to catch the bus, I already sort of commit to it by a hasty visit to the toilet. But if I used it the evening before, then I may even have forgotten my plan the next morning.

It is in these cases that one may feel that one had an intention long before one started execution, and these are of course the hardest cases for criminal courts to crack when they have to judge whether an attempt to execute a plan was initiated. These complications are however rare.

Normally speaking, if we want to represent statements about intentionality in relation to actions, we ascribe a plan: intended are all those events not interpreted as failures of the plan.

One cannot however simply say, that the agent “has a plan”. The term plan is used in too many different senses in knowledge engineering literature, which mostly considers the meaning of the terminology under consideration in relation to automated planning. There is a difference between the plan being executed by the agent, which is what we are interested in, plans under consideration by the agent, and “primitive” plans the agent knows of that are normally executable and that the agent uses as components for the composition of plans. Since mainstream planning literature calls all of these plans (cf. i.a. [58, 25]), or alternatively tasks in the case of *hierarchical task network* (HTN) planning (cf. [107]), it is not very helpful in differentiating between a plan being executed, and the plans manipulated in the mental process that resulted in an executable plan.

Here we distinguish between a plan, which is simply the mental object we manipulate when we are planning, and the task, which does seem to suggest commitment and is the thing we eventually execute. Tasks are executed, and the detailed specification of the task presumably requires planning before, and sometimes during, execution. Plans are more generally the structures being manipulated in planning.

The following breakdown, loosely based on the description given in [85], captures

the uses of intention-related terms in this book:

problem (objective, goal) is desired change

task is executed to solve a problem (achieve an objective, goal)

competence is to know a method for performing a task, which may involve decomposing it into subtasks

plan is a specification of a method for performing a task

action is executing a task, by causing events (occurrences of changes), resulting in success or failure

intention of an action is the task one executes: it simply refers to execution

The term competence is added because it is relevant to the legal theoretical concepts of (legal) competence or power.

A similar breakdown should in principle also be possible for designs (cf. section 2.2.4), which should be treated in analogy with plans. Also explicitly addressing designs in this chapter would however add little.

Most of the things in the list above belong to the mental realm, except for action, which bridges the gap between observable behaviour in the realm of the senses and the mental representation that gives rise to it:

Proposition 3. To describe behaviour as *action*, is to interpret it as the execution of a task:

$$\text{ACTION} \sqsubseteq \exists \text{executes.TASK}$$

Further elaboration on the structure of a task is not necessary for the purposes of this chapter, and quite hard to do without committing to a specific approach to automated planning. The action is an occurrent: it happens against the canvas of space and time. The action causes events, which are intentional if they are part of a successful performance of the task. Since the task specification is something that is mostly used in natural language to contrast actual performance with it, it makes sense to presume that there is some structural similarity between the action description and the task description.

The classification of intention as a relation between action and task may be contested. Some people may prefer to attach intentions to agents instead of actions, and to the (intended) result or *goal* of the task instead of to the task. These can however be interpreted as cases of metonymy (cf. with section 2.3). It is obvious to connect the intention to the action because of the temporal concurrence of holding the intention and executing the action. The concept action already presumes the deliberating agent. To say that an agent has an intention means nothing when separated from the action.

The preoccupation with goals is a similar case of metonymy. When one blows up a car with a bomb in the hope of killing its driver, one *intends* to place a bomb in a car, and to terminate the existence of a car, of a bomb, and of the life of a driver: all of these follow from the specified task, and will be recognized as intentional in law although the law may make additional distinctions. The term *goal* should be limited to the desired death of the driver that was an input to the planning process. The action has an intention, not a goal³.

³Note that for instance [14] also attacks the common conception that goal-directed and intentional

It is perhaps arguable that a more detailed representation of intention would demand its reification and connect it to action, task, goal, *and* agent.

The notion of action as task execution plays a central role in the representation of intent in *constitutive acts*, and in particular legal acts, as will be demonstrated in the next section.

The relation between actions and tasks however also plays a role in explaining certain other aspects of law. The most complicated ingredient for a fully developed theory of intentionality is perhaps the competence to perform a task: presumably this is based on some organization of knowledge which predisposes the agent to believing the performance of certain tasks will generally be successful. Moreover other agents are also able to ascribe competence to others in the context of task execution.

Examples of such “primitive” tasks that can be performed, or recognized, without further ado according to planning literature are such complex procedures as buying a quart of milk, painting a table, taking a plane to Melbourne, etc. To explain where these primitives come from I will appeal to Schank and Abelson’s *scripts* (cf. [246]), a concept developed for automated story understanding. Section 4.5 connects these scripts to *normality*, and discusses the relations between normality and normativity.

In law, even buying a quart of milk can however become a complicated affair, because it involves a transaction between two agents free to deviate from the script. If just action is already complex, transaction adds the complication of their interaction. On the level of transactions we find a new vocabulary – associated with Hohfeld (viz. [156]) who ironically thought of them as *fundamental* legal concepts – to gain traction over our problem. This new abstraction level is the subject of section 4.8.2.

4.4. Law as an Institution

Law is an institution, and therefore can be analyzed in terms of constitutive rules and institutional facts, backed up by an institutional ontology. When specifically applied to law we can speak of *legal rules* and *legal facts*, backed up by a *legal ontology*. The *legal act*, however, is an institutional act only because it was intended by the actor to be one: we cannot properly make the distinction between the constituting act and the constitutive act. Confusingly, theft is therefore merely a legal qualification of a natural act and not a legal act, but that some act constitutes theft *is* a legal fact, if we follow prevailing doctrine.

Proposition 4. An act that constitutes a *legal act* is constitutive: it is intended to constitute the legal act.

Example 4. The notion that one must intend to create a constitutional fact is captured by one of the following uses of the **constitutes** and **executes** property, applied to

are the same thing from another angle, pointing to a substantial body of research on *unintended but goal-directed behaviour*. This complication is however not taken into account here, since [14] considers goal-directed behaviour that would not be described as intentional.

an action:

$$\begin{aligned} \text{RAISEHAND} \sqcap \exists \text{executes.BID} &\sqsubseteq \exists \text{constitutes.BID} \\ \text{RAISEHAND} \sqcap \exists \text{executes.}(\text{RAISEHAND} \sqcap \exists \text{constitutes.BID}) & \\ &\sqsubseteq \exists \text{constitutes.BID} \end{aligned}$$

A RAISEHAND action, interpreted as the execution of a bid, constitutes a bid, or, better, a RAISEHAND action, interpreted as the execution of a RAISEHAND constituting a bid, constitutes a bid. Note that the term RAISEHAND is applied to both an action and a task, as is generally done in natural language. It is the use of the *executes* property that distinguishes whether we are talking about the real thing or its mental representation.

To position the role of intent in institutional reality I will distinguish (roughly following Sartor's account in [244]⁴) between legal rules that:

1. merely state that some event or situation constitutes an institutional fact,
2. state that some act intended to constitute some institutional fact constitutes that institutional fact,
3. are *intended* (by the legislator) to enable people to create institutional facts, and
4. are *intended* (by the legislator) to enable people to create institutional facts in order to pursue their own interests.

The distinction between the second and third type is tenuous: in some cases the legal rule only confers the benefit of legal recognition to actions that would also take place without it (for instance buying and selling), while in other cases the legislator creates a rule solely for the purpose of creating a recognizable way to achieve a certain legal effect (for instance a permit application procedure for gaining permission for constructing a shed in your garden). The legislator sometimes explicitly has to *create* a recognizable way of achieving certain novel legal effects.

Each category adds another constraint, and is as such a proper subset of the previous one. The classification is however purely by function: in essence they are all the same rules. Only the second type adds a discernable component: a use of the *executes* property. If we have to ascribe task execution to someone else, the *executes* constraint is however not going to add any information, since we will infer intention from the more readily observable aspects of the behaviour under scrutiny.

If the legislator intends to enable people to create some institutional fact, one usually speaks of (attribution of a) *power or competence*, and if the power or competence is attributed to help people pursue their *own* interests one speaks of a *potestative right* in Sartor's terminology in [244]. The link between pursuing one's own interests and the introduction of the concept of a right will be treated in more detail on section 4.8.

Pursuing one's own interests is however a problematic concept. It is easier to approach matters from the angle of pursuing someone else's interests. Civil servants for

⁴Note that this account however lacks the component of ascribed intention: intention is perhaps implicitly available in the distinction between productive and behavioural characterizations (cf. section 4.7.1) made there.

instance often have powers – linked to their role of civil servant – that they can only use as part of a predefined administrative task realizing an objective of the administrative body, and outside it they don't. Within the context of the task they have the power. In KBS this notion is rarely important, as we are usually dealing with routine tasks with a predefined structure.

Sartor discusses these rule variants as types of *normative conditionals*, and discusses the constitutive or “counts-as” rule as a subtype of them. Although I do not adopt this conceptualization, I do concur with his observation that all of them are simply typical variations of a single underlying pattern: that of a rule connecting antecedent to institutional fact. Sartor gives a more limited interpretation to constitutive and “counts-as” rules, which are in his opinion non-deontic of character. This distinction is problematic: the unlawful taking of a good that belongs to somebody else clearly *constitutes* theft, but the same formulation also – as a side effect – prohibits the unlawful taking of a good that belongs to somebody else if we understand the concept of theft to be *evaluative* in character, in this case being unambiguously negative.

Intuitively one would like to arrive at the conclusion that criminal law prohibits theft, but at least the Dutch, German, Italian, and English formulations do not explicitly *command* anyone not to commit theft. In section refsect:h4-norm and 4.6 I will discuss the relation between the constitutive and normative character of rules: here we only come to the conclusion that neither theft, nor the act that constitutes theft is a legal act.

Constitutive acts – in general – can be informal or formal, and by this we usually mean: backed by text. Text is taken to mean any representation preserved in a form whose existence is independent of both sender and receiver. Formal acts play a central role in law. Law uses formal constitutive acts to formalize constitutive rules, enact constitutive rules, and to repeal constitutive rules. We can therefore also speak of *formal rules*, constitutive rules created by a formal act, and in the case of the law these are fortunately explicit and unambiguously prior to their effects.

Definition 9. A *formal act* is an act of representation, preserved in a form whose existence is independent of both sender and receiver, of an institutional fact, with the intent to be constitutive of the represented institutional fact.

Example 5. In a literate society, the legislator often chooses for formal legal acts, since these make it easier to recognize intent (i.e. task performance). Instead of bidding by raising hands, auctions may also accept formal written or electronic bids:

$$\begin{aligned} \exists \text{result}. (\exists \text{represents}. \text{BID} \sqcap \exists \text{receiver}. \text{AUCTIONMASTER}) \sqcap \\ \exists \text{executes}. \text{BID} \sqsubseteq \exists \text{constitutes}. \text{BID} \end{aligned}$$

This means that an action that results in a message representing a bid to an auction master, with the intent to perform a bid, constitutes a bid. The result property is not the recommended way of representing what is brought about by actions, but strictly serves as a simple example.

Proposition 5. A *legislative act* is a formal act, that creates institutional rules, constitutive rules, and/or institutional facts required for correct interpretation or functioning of institutional rules or constitutive rules.

The legislative act is a strong argument for the institutional interpretation of law. A powerful criticism of the notion of the constitutive rule is to point out the suspect causal mechanism: if the rule backs the institutional fact, then obviously the rules exist prior to the institution, and if their results are recognized by an audience, then that audience should be able to verbalize the rules in effect.

Example 6. MacCormick in [192] discusses the example of queueing – a behaviour that a large part of the world population engages in – to show that the distinction between institutions and non-institutional conventions or normative orders are hard to draw using the concepts of constitutive rules. It is not obvious that the participants in a queue will be able to verbalize the rules in effect. What is clear, however, is that queueing behaviour will be increasingly formalized – in the form of a visible and recognizable queue – as more people start to participate in it and queue-jumping starts to occur.

While we can doubt the construct validity of the constitutive rule in general, we can rest assured that the analysis of institutions should at least work for the law already in operation, at least to the extent that it formalizes its constitutive rules. Only constitutional law itself, which gives the rules for creating and changing the rules, usually came about in a rather messy fashion (cf. for instance [260]).

Law of course is not a monolithic system. There are many institutions that assume legislative power, and there is no guarantee that these institutions cooperate in harmony in creating their legal ontologies, legal facts, and constitutive rules, or even that they recognize each other. A legal institution can selectively recognize the institutional facts of another institution by way of constitutive rules, which may be formal or informal. It can also adopt the ontology of another institution without recognizing its institutional facts altogether: they merely use the same terminology, but ignore each other's institutional facts completely. Clearly marking off the boundaries between different institutions, or deciding what it is that is being recognized, is not always trivial.

Proposition 6. A legal institution can recognize the institutional facts of another institution by way of constitutive rules.

Example 7. A simple example is the recognition of driver's licenses. The law in many countries regulates who is allowed to drive, in order to allow only people who are competent to drive on public roads. When the legal system finds someone competent to drive it issues a drivers license. This (constitutive) act is intended to create the institutional fact that this person is licensed to drive, and it is formal: it results in a physical representation of this institutional fact in the form of the drivers license. The drivers license is *prima facie* evidence for the license to drive, but the license to drive may be retracted (constitutive act), in which case the drivers license becomes void if this happens in his absence.

A country may recognize the institutional fact of another country by way of a constitutive rule, or it may simply directly recognize the foreign physical driver's license as proof of competence by way of a constitutive rule. The two mechanisms have different results in relation to voided licenses, but to find out that the driver's license is voided the two countries obviously need to exchange institutional information.

$$\begin{array}{c} \text{BELGIUM:DRIVINGCOMPETENCE} \\ \sqsubseteq \exists \text{constitutes.NETHERLANDS:DRIVINGCOMPETENCE} \end{array}$$

$$\begin{array}{c} \text{DRIVERSLICENSE} \sqcap \exists \text{represents.BELGIUM:DRIVINGCOMPETENCE} \sqsubseteq \\ \exists \text{constitutes.NETHERLANDS:DRIVINGCOMPETENCE} \end{array}$$

In the first example it is an institutional proposition, belonging to another institution, which is recognized. In the second example a physical object is recognized.

Institutions may use very similar, or even exactly the same, institutional ontology without sharing constitutive rules. The similarity is in this case very shallow, as there is no link at all between both institutional realities. The institutions are in effect sharing the same abstract models.

4.4.1. Constitutive Rules and Knowledge Representation

At this point we have the basic ingredients for representing legal rules as constitutive and institutional rules: the **constitutes** and **executes** properties, relating disjoint ontological strata.

It is however incorrect to say that the domain and range of the **constitutes** property are disjoint in general; They are disjoint from the perspective of any one specific institution. Institutions do however recognize each other's institutional facts as brute facts. The institutional interpretation is an abstract model, to be implemented for any specific institution i by creating a specific property $i:\text{constitutes}$ that does relate disjoint domains.

There is also another important point to be made here: Although the logical form of institutional rules and constitutive rules is being discussed in this chapter, we have not fully committed to the notion that legal rules *are* institutional rules and constitutive rules. As we will see in section 5.2.1, they are not.

The simple examples used so far in this section have all used OWL DL axioms instead of the defeasible rules introduced in section 3.4.2. The question is however whether the use of constitutive rules should be represented with OWL's monotonic \sqsubseteq operator or as a kind of ampliative reasoning. *Institutional* rules define the institution's ontology and are properly modeled by the monotonic \sqsubseteq operator. *Constitutive* rules however perform a mapping from one ontological stratum to another ontological stratum and this mapping may be defeasible.

Each of these strata have their own space of possible models as described by their ontology, and there is no guarantee that constitutive rules always perform an unambiguous mapping from one stratum into the other one. It is in fact not sensible to

assume that they do, and we need to resort to the defeasible rules of section 3.4.2 if they don't.

Constitutive rules may come in the $C_b \sqsubseteq \exists \text{law:constitutes}.C_i$ *indicator* and the $\exists \text{law:constitutes}.C_i \sqsubseteq C_b$ *requirement* varieties, or the combination of both, where C_b is a concept from brute reality, and C_i a concept from institutional reality. To reduce maintenance we will assume that it is the $C_b \sqsubseteq \exists \text{law:constitutes}.C_i$ that is defeasible, and should be modeled by a rule of the following form:

$$\begin{aligned} &(\text{DEFAULT } r \text{ (known}.C_b) \text{ (free.}\exists \text{law:constitutes}.C_i) \\ &\quad \text{(assume.}\exists \text{law:constitutes}.C_i)) \end{aligned}$$

A matching brute reality and legal reality is not guaranteed by the model set described by the source of law: it describes the part of brute reality that specifies the interface to the institution, not the part of it that doesn't. The purpose of the rules is not to describe the ontology of the brute reality the institution is deposited on top of. The ontological coherence of institutional reality and brute reality take precedence over the applicability of constitutive rules.

Proposition 7. A *legal fact* can only come into existence if it is *consistent* with the settled facts, and if there is some constitutive rule that *indicates* it.

Consider the following detailed, but simplified, real world example, loosely based on [244], who models a similar structure in two different rules:

Example 8. Let us say that 1) the taking of something constitutes theft, and that 2) theft must have been performed with the intention to appropriate. Interpret this as a *set* of two rules, in no particular order.

The separation into two rules presumably comes from Italian criminal law; The German and Dutch version both have a “with the intent to” sentence fragment *embedded* as a condition in their variant of rule 1, which changes matters. The issue is not whether these rules accurately interpret Italian law, or any other law, or the intended reading in [244].

The freedom of interpretation increases if rule 2 reads: *The* theft must have been performed with the intention to appropriate. To come to this interpretation for rule 1 and 2, *the* in rule 2 must refer to a theft as in rule 1. See section 5.3.4 for this alternative reading, which depends on the structure of the writing that expresses the rules. In this case rule 1 and 2 are really one rule, just as would be the case if 2 was embedded in 1 as a sentence fragment. Another variation on the same theme is the following formulation: *The theft of rule 1* must have been performed with the intention to appropriate. If the two rules are however structurally separate entities, and there is no *the* connecting them on the linguistic level, they should not have to be joined in a rule or OWL class.

In the rules as I phrased them the (observable) taking *indicates* in my view the possible existence of a theft, while the intent to appropriate (which happens to be not observable) is a *requirement* to classifying something as theft.

The sense of the verb *must* as used here hints at a necessary condition, or the “epistemic obligation” to apply the rule. [244] gives another interpretation of the

verb *must* used in this sense. The verb is in my view however a red herring: the interpretation of the sentence does not change if it states that theft is performed with the intention to appropriate. What matters is whether or not the sentence can be interpreted as being embedded in rule 1.

Observe that this example illustrates that although we are usually “deriving” institutional reality (the theft) from brute reality (the taking and the intention with which it is done) arrows may well point the wrong way in rules:

$$\begin{aligned} \text{TAKING} &\sqsubseteq \exists \text{law:constitutes.LAW:THEFT} \\ \exists \text{law:constitutes.LAW:THEFT} &\sqsubseteq \exists \text{executes.APPROPRIATION} \end{aligned}$$

The first rule is a *prima facie* sufficient condition for theft, and the second one a necessary one (see figure 4.1).

This interpretation is as much saying that takings always happen with the intent to appropriate. If you take something from the shelf in the supermarket, you do not have the intent to appropriate it: there is no acceptable mapping from brute reality to legal reality if we interpret the constitutive rules as OWL axioms. This is however a perfectly reasonable set of rules.

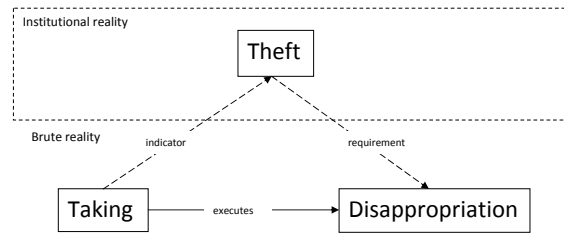


Figure 4.1. The act of taking functions as an indicator of theft, among other possible indicators. The intent to appropriate is however a *necessary* condition. The total of rules describing the relation between a theft and the brute facts underlying it define its constituting base.

Consider what happens if in the future we would have to add 1b) illegal copying (also) constitutes theft⁵:

$$\text{COPYING} \sqsubseteq \exists \text{law:constitutes.LAW:THEFT}$$

This creates one of the major maintenance problems often encountered in the representation of sources of law. Rule 2 should in this phrasing have an effect on thefts indicated by rule 1b, and therefore cannot be embedded in rule 1. Moreover if rule 2 reads *the* theft [...] the article *the* suddenly resolves to rules 1b if it is inserted between rule 1 and 2.

⁵This is an ontologically absurd extension that will never happen, despite the clamoring in the media about copyright violation as theft, but the descriptive fidelity of the example is not the issue. The example is about the two rules as given here, approached without assumptions about how things should work.

An embedding consists of combining the rules into one by making the necessary condition part of the sufficient condition:

$$\text{TAKING} \sqcap \exists \text{executes.APPROPRIATION} \sqsubseteq \exists \text{law:constitutes.LAW:THEFT}$$

The appropriation is in this case not a necessary condition. This is a very common approach, and the solution chosen by [244], who clearly has an embedding in mind: theft is the conclusion of the rule. The rule now settles for an *unknown* intent to appropriate in OWL DL, since we may simply assert that something constitutes a theft without restriction. The problem is also that the rule has to be changed to accommodate rule 3 about COPYING.

From a database-oriented point of view one may also argue, quite plausibly, that constitutive rules only map from legal reality to the brute one, and take the position that all constitutive rules therefore specify necessary conditions:

$$\exists \text{law:constitutes.LAW:THEFT} \sqsubseteq \text{TAKING}$$

$$\exists \text{law:constitutes.LAW:THEFT} \sqsubseteq \exists \text{executes.APPROPRIATION}$$

If we however add rule 3 theft would require taking a good *and* copying information instead of *or*, which can't be the intended interpretation, so we still have to modify an OWL axiom to accommodate a new rule.

Reformulating the rules to achieve specific effects as in the example is not a good idea from a maintenance point of view. Rule 1 and 2 should only be taken together in knowledge representation if one is of the opinion that they *express* only one rule. In that case, which may be indicated by the use of the article *the*, rule 1 and 2, when read separately, are clearly meaningless. Since there is no need for isomorphism with rule 1 and rule 2 respectively, there is also no need for a special type of rule insertion an extra condition in other rules, as [244] argues.

Both reformulations above try to answer the question which thesis is defeated, taking into account ontological stratification: the first one presumes a settled brute reality and explains it in terms of legal facts, while the second one assumes a context in which one is explaining a settled legal fact in terms of a matching brute reality.

These approaches are both correct interpretations of the use of stratification: one takes one stratum as the independent variable and the other one as the dependent variable. But the source of law should not be directly represented in these ways: the theory put forward by the source of law is after all perfectly consistent in itself.

The set of rules constructed here does not make that possible. The solution for this set of rules, based on the assumption that legal facts only come into existence if they are consistent with the settled (brute *and* institutional) facts, is to guard the derivation of the legal fact with the requirement that this derivation should be consistent:

$$\begin{aligned} &(\text{DEFAULT } r_1 \text{ (known.TAKING)})(\text{free.}\exists \text{law:constitutes.LAW:THEFT}) \\ &\quad (\text{assume.}\exists \text{law:constitutes.LAW:THEFT})) \\ &\exists \text{law:constitutes.LAW:THEFT} \sqsubseteq \exists \text{executes.APPROPRIATION} \end{aligned}$$

The derivation of the institutional fact of a theft is based on the *ampliative* inference from an observed taking and the non-ampliative inference to the intent of the taking. The two rules in other words describe a typical *generate & test* approach (cf. section 2.4.1). Note that logically speaking, if one makes the open world assumption, it is possible to assert that something is a theft without proposing that that something is a taking.

The reason we don't do that is because we view the law as a category of systems that can only be manipulated through the interface of socially recognized constitutive rules. For legal facts this means that it is relevant which rule was applied (which is addressed in section 5.2.2).

In rare cases we may also want to use constraints.

For purposes of application in a KBS one also needs to decide whether one assigns a burden of proof to the user with this set of rules. In this case one might for instance want a separate argument (as defined in 3.3) for the ascription of intent instead of allowing intent to be justified by default by the taking itself, since the intent is one of the constituting facts:

$$(\text{CONSTRAINT } r_2 (\text{known}.\exists \text{law:constitutes.LAW:THEFT}) \\ (\text{fails}.\exists \text{executes.APPROPRIATION}))$$

Even though $\exists \text{executes.APPROPRIATION}$ is a dispositional belief in a knowledge base using the necessary condition stated above, there is no assertion that refers to the task being executed by a constant. Note that also this constraint does not break the isomorphism requirement: it clearly refers to rule 2 only.

These constraints are not directly relevant for the reusable representation of the source of law: one does not automatically assume the burden of proof just because one applies the rule to explain people's behaviour. Burdens of proof must be met by formal justifications, not by reasoning in general. Datalog-style constraints are, as opposed to OWL's *tbody* axioms, defeasible. The falsification of a constraint is commonly used to drive interaction with a user.

Proposition 8. A *burden of proof* can be represented with a constraint.

The representation of burdens of proof by constraints is sufficient for simple Datalog-based KBS as sketched in section 3.4.1, where one can simply pick a set of constraints on the amount of elaboration necessary for acceptable arguments that one would like to enforce in the (static) KBS problem setting.

If one wants to make the applicability of burdens of proof *conditional* on features of the problem solving setting, a more advanced representation is necessary. Explicit reasoning on the allocation of burdens of proof is for instance often required in legal argumentation (cf. for instance generally [229]).

4.5. Norms, Normality and Normativity

In the preceding sections the institutional interpretation of legal knowledge was developed. There is however another important, in many accounts even more important,

aspect to the law. The law is often said to formalize *norms*. In this section the notion of norm is investigated, from a cognitive perspective, before we turn to the normative rule of section 4.6. This section also sets the background for chapter 6, which will address the *normative order*.

We commonly make a distinction between *normality* and *normativity*. Normality refers to the norm from a descriptive perspective: it merely describes an observed regularity in the behaviour of instruments or agents. Normativity only applies to the behaviour of agents, and communicates something in addition to the mere regularity of behaviour: the behaviour is motivated by the recognition of the norm by those whose behaviour is consistent with it.

Normality and normativity are different things, and obviously should be treated as such. We learned long ago from Hume in [160] that “ought” does not follow from “is”. This relates to the discussion in section 2.2.2 about the truth value of norms: no constellations of facts can ever prove the truth of a norm. The norm can be communicated between agents, and it can be recognized and adopted, but it is never the spontaneous result of logical reasoning from facts.

There are however a number of ways in which normality and normativity, different sides of a same coin, interact. In the commonsense conception of justice this relation is clear, and fairly well-understood. Montaigne already pointed this out in [90]:

I am prepared to forgive our own people for having no other model or rule of perfection but their own manners and behaviour, for it is a common failing not only of the mob but of virtually all men to set their sights within the limitations of the customs into which they were born.

The very word morality derives from *mos* or *mores* (in plural) *maiorum*, the ways of our ancestors. That custom is the basis for what is moral seems to involve a shift from what is to what ought to be.

Normativity as a source of behaviour is the traditional province of Computer Science & Law and Legal Knowledge Engineering. Normality is the province of descriptive sciences, and in Knowledge Engineering is implicit in notions like *defaults* (cf. generally [175]) in reasoning, world knowledge and *common sense knowledge* (cf. [63]), or Schank’s scripts (cf. [246]).

Not all norms have something to do with behaviour of intentional agents. The *norm* is identified by its *epistemic role* in problem solving and not something that exclusively *belongs* to the vocabulary of the legal domain, or to that of morality. A norm is a standard of performance, a measurement scale. It is used to predict performance and to diagnose performance contrary to expectations, regardless of domain (cf. [266]).

The norm is conceptualized differently in different contexts: When assessing readings from a broken instrument – for instance some circuit board – a norm is for instance the distribution of scores obtained from a correctly functioning norm group, or some margin around the specified ideal functional mapping from input to output that describes the expected and intended behaviour of the instrument. The expected behaviour of the instrument is also intended: the instrument was designed by an intentional agents with some intended purpose in mind. This process shares little with law except the epistemological roles (assessment, norms, case, qualification) in the problem solving process.

In sociology, a *social norm* is a pattern of behaviour expected of an agent within a particular society in a given situation. In this context the norm is interpreted as directly reflecting some preference of the involved agent; Given that an agent is aware of a number of behavioural alternatives and made a choice between them, we can infer that the agent revealed a *preference* since choice and preference are interdefinable. If we observe a pattern of agents revealing the same preference and suspect that there is some social mechanism that explains how that shared preference arose, we may infer that there is a social norm that explains that preference. What makes the behaviour normal is simply the fact that it is *expected*.

There is a presumption, for instance in predicting the consequences of legislation, that people will normally do what is normative. In other words, it is normal to comply with the norms.

In legal theory we also find mention of custom as a “source of law”. The unlawfulness of taking a good (theft) is for instance judged against a cultural context. Is it normal to take something from a shelf in a supermarket? Is it normal to take your neighbour’s ladder to get your cat out of a nearby tree? The assumption is generally that what is normal is lawful, and legislation prohibiting what is normal will be less easily accepted.

Not recognizing this can lead to serious overestimation by legislators of their power to change society. A legislator that likes to be in control should limit himself to instructing people to do what they are willing to do.

Normality and normativity are just two ways of explaining patterns in behaviour: something is normal according to the observer because the observed pattern happens to reveal it to the observer, and it is normative if it motivates (or “causes”) the choices that lead to the pattern. Normality appears as a social fact and a social value, because the norm is used as a standard of reference against which deviations from normality are evaluated. The normal has both factual and normative force. Associated with the former is custom which is revealed through observation as a matter of fact. Associated with normativity are morality, ethics, and the law.

To explain how normality and normativity relate, and what effects this relation has that are relevant for representing and evaluating legislation, I will introduce the concept of *status quo bias*. It is central to the formation of common sense standards of justice or fairness, and can for instance be taken into account in comparison and evaluation of legislation on behalf of the legislator to judge the effectiveness and efficiency of legislation.

This concept is related to a version of economic theory called *prospect theory* (as opposed to *utility theory*, which is almost universally considered a standard for rational decision-making), credited on Kahneman et al. in [164]. His account, based on results from experimental psychology, revolves around the *reference transaction*, and its role as a yardstick for what is fair. That is, what we become accustomed to (the status quo) attains a normative status against which deviations to our disadvantage are considered unjust.

Status quo bias refers to the finding that in games an option is more desirable if it is the status quo for no other reason than that it is so designated (cf. [165]). Status quo bias is often discussed in relation to the endowment effect and loss aversion (cf.

[166]). Status quo bias differs from the other two in that it does not (even) depend on framing changes in terms of (possible) losses and gains.

The endowment effect refers to the differential weight placed on the value of an alternative depending on whether one “possesses” the alternative and is faced with its loss or whether one does not possess it and has the potential to gain it. Losing a given alternative, which is part of one’s endowment, is felt to be a greater loss than the corresponding chance of gaining it when it is not part of one’s endowment. Losses are more heavily weighted than foregone gains.

Furthermore, a certain degree of inertia is introduced into a choice process since things that are already included in the individuals endowment will be more highly valued than those not held in the endowment. Kahneman and Tversky introduce the term loss aversion to capture the relative steepness of the loss-portion of this value function relative to that on the gain side of it.

Kahneman et al. in [164] developed the idea of a *reference transaction*: a relevant precedent that is characterized by a reference price or wage, and by a positive reference profit. In the course of life, certain practices and patterns establish themselves and attain the status of normality. If someone disrupts this normal course of affairs, then the disruption will be considered unfair if:

- it is to the advantage of the other and to your disadvantage, and
- if there is insufficient justification for the disruption.

The involved parties are entitled to the terms of the reference transaction. The reference transaction, as a custom, is part of one’s endowment.

Example 9. Consider the following example from [164]: A hardware store has been selling snow shovels for \$15. The morning after a large snowstorm, the store raises the price to \$20. The reference transaction, given by precedent, is marked by a price of \$15. The price increase transgresses the reference transaction. In a telephone survey, 82% of the 107 respondents deemed the price increase to be unfair.

Kahneman et al. (cf. [164]) claim that individuals consider the reference transaction to be fair *because* it is normal: individuals do not have any anterior criterion of justice which informs their judgments of fairness. Kahneman et al. even suggest that any stable state of affairs tends to become accepted as fair eventually. The reference transaction consists in what people have become accustomed to but it also becomes a normative yardstick for assessing the fairness of deviations from itself. This is perhaps philosophically suspect from Hume’s point of view at least, but has also been recognized as an important factor in the sociology of law (cf. for instance [119]).

Kahneman’s work in psychology has had quite an impact on economics, eventually resulting in a Nobel Prize for economics in 2002, because it attacks the fundamentals of utility theory as a predictor of actual human behaviour in a systematic and empirically well-founded way.

The notion of a reference transaction is naturally connected to the *script* in knowledge representation, and my earlier tentative specification of what competence is: to know a method, gleaned from episodic memory, to perform a task. If the snow shovel is made significantly more expensive, or the quart of milk is no longer for sale, etc.,

this is unfair because it takes away something from us: the potential to do certain things following a predictable script.

The champion of scripts in knowledge engineering is Roger Schank (see [246]). The key idea involved in scripts is that our knowledge of concepts, events, and situations is organized around expectations of key features of those situations. Many choices are made in the context of a script: you do the same thing every time as long as the salient features of the scripted situations don't change.

In Schank's theory, memory is episodic, i.e. organized around personal experiences rather than semantic categories. Generalized episodes are called scripts. Schank's ideas were developed in the context of story understanding, but have also played a key role in the development of hierarchical and case-based planning algorithms (the elusive store of *skeletal plans* and *primitive plans*; viz. [107]). Scripts are also organized around the *social roles* (buyer, employee, customer, student, citizen, father, etc.) one plays in life (see for instance [198] for more on social roles).

Proposition 9. Whether behaviour is *normal* is assessed against scripts associated with the social role one assumes.

It only makes sense that institutional settings also get internalized as scripts, as reference transactions, i.e. become the normal way of doing things, and a guide for predicting the behaviour of others.

The legal *normative rule* is a constitutive rule with the purpose of creating normative order. It does so by way of positive or negative qualifications of behaviour. The unlawful taking of a good that belongs to somebody else *constitutes* a theft, theft negatively qualifies the taking, and gives license (through other rules) to other agents to do things to the thief that are generally speaking not in the thief's interest. This regulative function of a fraction of legal rules is however not the only way in which the law promotes normative order.

Both normative rules and other constitutive rules of the institution of law tend to affect what is perceived as normal. Legislation, in so far as it induces certain types of behaviour, has an impact on habit formation and hence on the formation of the standards of justice of the future. If the legislator sets a standard for legally recognized sales transactions, actual sales transactions will tend to conform to the prescribed model as it gradually becomes the new reference for sales transactions.

Normative rules qualifying undesirable behaviour in sales will be phrased in terms of a deviation of these reference transaction: in the interest of predictability it is therefore a good idea to follow that script, and if someone deviates from the script *to your disadvantage and their advantage*, it is only natural to presume that what he does must be illegal.

Proposition 10. All law, not just its normative part, promotes normative order insofar as it has an impact on habit formation and hence on the formation of the standards of justice of the future.

Law has a *de facto* effect on prevailing social standards through habit formation and the formation of standards of justice. This is however not all. The legislator obviously also has *intentions* with the normative order that he tries to realize through legislative action.

4.5.1. Legislative Action

As an observation about legislative action the reference transaction is very relevant. Firstly, it predicts, as noted, that the addressees of changed legislation will generally speaking cling more fiercely to existing reference transactions than utility theory would predict.

Secondly, it gives a basis for understanding *norms of analysis* as understood by Hettich and Winer: the norms used to evaluate proposals for legislation (cf. [146, 145]), which will be addressed again in section 6.7.3.

We for instance expect the following from the combination of Schank's scripted behaviour in planning and the formation of standards of justice around Kahneman's reference transactions:

- Stakeholders will usually evaluate a change in legislation by considering roles – buyer, employee, employer, etc. – they typically or regularly play in reference transactions, and not their future position as a rational agent that could change roles. The buyer of a snow shovel will not contemplate becoming a seller of snow shovels. Political lobby organizations that try to influence legislation obviously do so from the conception of a role the members they represent play.
- The legislator will also usually take into account the reference transactions created by the existing legislation and evaluate the new one as a deviation from it. It is possible to extrapolate roles in these reference transactions to populations filling the role, certainly in macro-economic modeling of taxation and social security⁶. The estimated number of people playing each role will often be considered constant in the legislator's evaluation, except when “social movement” from role to role is the purpose or a salient aspect of the change.
- It is easier to change non-salient aspects of the reference transaction than salient ones. Making people pay more often by making the transaction occur more often for instance works better than making them pay more in the transaction.
- Peacock [221] makes a more ominous prediction about the legislative process: the reversal of policy measures is more difficult than their introduction, particularly if the losers of a reversal are clearly identifiable. Consequently, according to [221], a government should consider this effect of its policy measures before implementing them. If a policy is intended as a short-term measure, it may, when it comes to reversing it, have become so customary that reversal is very difficult or nearly impossible.

The appropriate normative yardsticks for evaluation of legislation measure deviations from abstract reference transactions. The norms and concepts by which a new version of legislation is going to be judged are probably largely based in the existing version. Changes made will be Pareto improvements – at least one individual better off, without making any other individual worse off – but in terms of *social roles* identified in the relevant reference transactions and not of individuals, which would be an impossibly hard problem. In [263], Tanghe also makes this distinction in political

⁶Dutch readers will be familiar with the continually updated *koopkrachtplaatje* (roughly *purchasing power table*) and the population segment they are in.

discourse focusing on inequality and its relation to injustice: this concept in its most widely accepted sense revolves around equality of ability to assume social roles, Tocqueville's *equality of conditions*, but this form of equality has arbitrary results when extrapolated to specific populations of one's choice.

These social roles, and the expectations about them, play a central role in organizing the knowledge we use in planning.

4.5.2. Representing Normal Behaviour

Kahneman's reference transactions directly relate to the perceived competence to perform mundane or "primitive" tasks. The effect of constitutive rules is to create a script through which one can achieve certain legal effects, i.e. recognized by others. If these effects are believed to be beneficial by the agent, the agent will generally speaking use the script in planning. Conversely, other agents will ascribe plans to the agent based on the same scripts and assumptions about his beliefs and desires, and recognize the legal effects he brings about. Normal behaviour is simply expected behaviour.

If legal effects of constitutive rules are believed to be detrimental by the agent, the agent will use the script in planning to recognize bad plans. The most important rule of this kind is obviously the normative rule. Agents will however also evade for instance liability to taxation, or legal effects that create costly duties, liabilities, or disabilities.

Example 10. For instance: if you need a building permit for a shed higher than 2.5m, this is an argument *against designing such a shed*. Another nice example comes from the CLIME project (cf. [284]): There is a maximum size constraint for ships allowed entry into the Panama canal called PANAMAX. This rule is however irrelevant for operators of ships: ship designers all over the world use it as a hard constraint for ship design.

We apply the notion of "normal" behaviour both to objects and to other people (and perhaps, following [95], to ourselves). Certain categories of objects, tools, machines, artifacts generally, have a *function* or instrumental role. We have designed them for the purpose of exhibiting certain predictable and useful behaviours when we interact with them in certain ways, and this allows us to perform certain tasks. We also classify them by that function. If their behaviour deviates from the expected one, usually in a detrimental way, they are broken and need to be repaired or replaced.

When interacting with other people we do exactly the same. To buy a quart of milk, you have to find someone who can fill the agent role of *seller* of milk, etc. We judge deviations from expected behaviour in the same way, except that the remedies are of a different character. [198, 153] point out that social roles are identified by the set of actions that can be performed in that role. The social role is also a separate object whose existence depends on the "brute" object in the stratum below. This is for instance shown by the application of terms like "good" to roles like seller, student, or cook: the properties apply to the agent in the role, but not to the agent separately (cf. [139]). Legal roles are simply social roles in a legal institution. Roles are functions when applied to objects, and agents when applied to persons as the active and efficient

cause of events. Agent roles that depend on recognition, and are therefore institutional, are social roles, and if they depend on recognition of the law they are legal.

These sets of actions connected to agent roles can be composed into scripts with interacting participants, who each take a role. To recognize a series of events as an instance of the script, is to expect that the other agents will act according to their role. If they deviate from the script, a new explanation of the situation is needed. Executing a task is to play your role in a suitable script you selected to solve a problem or achieve a goal.

To adopt the agent role in a script one must take a number of normal decisions, expressing a set of preferences, which are attributed by others to the agent observed to be involved in the script. Deviation from the scripted conditions leads to a re-evaluation of one's behaviour.

For the representation of sources of law, this information about the planning process has only an indirect relevance. As stated in section 4.3 we do not commit to a specific account of planning, and therefore have no model of what this information structure looks like that relates attributed social roles to expectations about the intentions of the filler of the social role.

Since our knowledge about normal behaviour relates to the agent role rather than the person, we should however explicitly distinguish the person who plays an agent role from the agent role itself. To be an agent (role), is to be played by a person and to be the actor in an action:

$$\text{AGENT} \equiv \exists \text{playedBy.PERSON} \sqcap \exists \text{actorIn.ACTION}$$

The higher order “set of actions” *belonging* to an agent role is simply indicated by the following type of axioms (the actor in a theft is a thief, and vice versa):

$$\exists \text{actorIn.LAW:THEFT} \sqsubseteq \text{THIEF}$$

The notion of agent roles is essential to the interpretation of competence, including legal competence, and normative order. The term *agent* in this book, and specifically in section 4.7, does not equate to person.

Interestingly, however, the normative rule is often explained without any appeal to agent roles.

4.6. Legal Normative Rules

There is a simple criterion to distinguish *normative* constitutive rules from other constitutive rules: the normative rule can either be violated or it can cancel the violation of another normative rule. The normative rule can be defined in terms of the institutional legal fact it can create: the *violation* of a normative rule, or the cancellation of the violation of a normative rule. The first one could be naively represented as:

$$\text{EX:SOMETHING} \sqsubseteq \exists \text{law:constitutes.LAW:VIOLATION}$$

As already indicated in section 4.1, normative rules come in the three major flavours of obligation, prohibition, and permission. The significance of the normative rule of

the obligation or prohibition type in legislation is in its effect as a constitutive rule: by stating that one ought to do or bring about x , it states (as a matter of terminology) that not doing or bringing about x is a violation.

It is the institutional qualification violation, which is inherently *evaluative* in character, that opens up the interpretation of the act as wrong in the deontological sense. It is this interpretation that makes *deontic logic* an obvious tool for representation of law. The effect of a permission is to cancel what would otherwise be a violation: it only means something in the presence of other, conflicting obligations or prohibitions.

Proposition 11. A *normative rule* is a constitutive rule that derives its violation, or the cancellation of a violation, from one or more constituting facts.

The institutional ontology of law incorporates inherently evaluative concepts like violation, obligation, crime, theft, etc. because it is intended to evoke the deontological interpretation, just like the institutional ontology of chess appeals to concepts like winning and losing because it is intended to evoke competition, and a manufacturer of hammers calls its products hammers because it intends to evoke an evaluation of their usefulness for delivering blows, which would not be conveyed by calling the hammers paper weights.

The purpose of law is to institutionalize normative order.

The concept of violation also evokes the concept of *punishment*. Not because violations are always followed by punishments, but because punishments *are by definition preceded by* violations. The preceding violation is the terminological *conditio sine qua non* of punishment⁷ Punishments are obviously not just disagreeable consequences of one's actions, but intentional and predictable counteractions by others intended to suppress violations.

Moreover, as already suggested in section 4.2, the following should also be considered (inter alia) a normative rule: "The taking of a good, that wholly or partially belongs to another, with the intent to unlawfully appropriate it, constitutes theft and will be punished with a prison sentence of at most four years". It is not the verb used or any other structural regularity in the expression that identifies the normative rule, but its inherent evaluativeness. It works as a standard constitutive rule for determining a legal fact of theft, negatively qualifies it by using a term – *theft* – that's inherently negatively charged⁸, and concludes with an institutional rule capping the length of prison sentences for an act that constitutes theft⁹.

The reference to *punishment* for theft makes this immediately recognizable as a normative rule. Not stealing is affirmed as a norm by the normative rule.

The classification of a rule as normative immediately puts a rich vocabulary of normative positions at our disposal. As already pointed out, normative rules are most commonly phrased as obligations, prohibitions, or permissions, but the surface form

⁷[141]: "Legal responsibility is the liability of a person to be punished, forced to compensate, or otherwise to be subjected to a sanction by the law."

⁸See for instance [139] for more about the relation between concepts and value statements.

⁹To represent this last rule transparently, one would first need to fix an interpretation of the more general rules of criminal procedure.

of the rule is – contrary to the examples given of constitutive rules – not determinative of its interpretation.

The ontology of normative positions is shared among legal institutions, and is indeed an identifying feature of legal institutions. This merits special attention for this ontology. Since this ontology is however based on a abstraction level beyond that of (institutional) legal rules created and communicated by sources of law, the formal account of normative rules is treated separately in chapter 6. This ontology is not found *in* the sources of law, but is rather background knowledge to the interpretation of the (intended or expected) effect of the source of law on behaviour.

In discourse we encounter the use of normative notions like obligation both in a generic sense, as an identification of a type of surface form of rules, and as a normative position in context relativized to specific objects. Consider the following two examples:

1. If (seller) offers to sell some (good) to (buyer), and (buyer) agrees to buy the good then (seller) has an obligation to sell the (good) to (buyer).
2. I have an obligation towards John, to sell him the painting I offered him for sale *if* he agrees to buy the painting.

The first one is a rule. The second one is partially filled in script based on the rule that describes a legal position: it indicates that some of the conditions for deriving a violation of the rule have already been met, and that if John agrees to buy the painting, which is outside my control, not selling the painting to him would constitute a violation. One is in a “state of obligation”. Clearly, the obligation and the rule are separate entities: the rule apparently gives rise to states of obligation, which are also institutional facts:

Proposition 12. A *normative rule* can be used to derive a state of obligation to do something, or the cancellation of such a state of obligation, from one or more facts.

The state of obligation is clearly ontologically related to the institutional fact of violation. Obligation is traditionally represented by an operator O , “it ought to be that”.

Depending on the kind of logic and the type of problem considered, knowledge engineers choose for different conceptualizations of normative rules.

There are several simple rule-based approaches that try to approximate norms in Datalog rules with notions like violation or obligation (cf. for instance [162]) wherever necessary.

A principled approach to representing normative rules as rules that derive an institutional fact from brute facts is the functional approach of Valente in [266]. It depends on a value function v mapping from descriptions to the ordered set $\{allowed, silent, disallowed\}$, where violation obviously maps to disallowed, and allowed cancels a disallowed in certain circumstances:

$$v(c) = \begin{cases} \textit{disallowed} & : O(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{disallowed} & : F(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : O(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : F(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{allowed} & : P(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : P(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{silent} & : \textit{otherwise} \end{cases}$$

$c \in \mid \phi \mid$ means: c is a member of the extension of ϕ (i.e. $\mid \phi \mid$ is the set of worlds w such that $M, w \models \phi$). This representation is clearly committed to considering normative rules as constitutive rules that derive their own violation (here disallowed), or the cancellation of the violation (here allowed) of another norm. Silence is the *absence* of an institutional fact.

Its strength is in dealing with permission and handling conflict between norms (cf. for instance [286], and section 6.2). A weakness is that it cannot be used to detect a *contrary-to-duty obligation* (cf. section 6.4), which is one of the tests we will use in chapter 6. [266] is satisfied with the observation that this approach does not lead to the contrary-to-duty paradox.

Valente's account strongly influences the view of normative rules in this book, in which we will adopt the notion of allowed and disallowed as institutional facts. Moreover, the institutional fact only relates to the specific normative rule. This lead us to a refinement of proposition 11.

Definition 10. A *normative rule* n is a constitutive rule that derives $\textit{disallows}(n, i)$ or $\textit{allows}(n, i)$ from constituting facts about i .

Logical accounts of *obligation* explain certain properties of a family of concepts revolving around *subjunctive betterness* (cf. [89, 195]): they mainly express that certain entities *would be better* than others. It is this account that is relevant to planning and to plan recognition from observed behaviour.

Very well-developed in computer science & law literature is the modal *deontic logic* approach, which turns obligation, prohibition, and permission into modal operators and works from the observation that obligation and prohibition are interdefinable ($O\alpha \equiv F\neg\alpha$), and obligation implies permission ($O\alpha \rightarrow P\alpha$). What deontic logics are especially good for is capturing the uses of obligation in a positional sense, as explained above.

The modal interpretation of O is a nonempty model set of deontic *alternatives* in such a way that what ought to be is the case in its deontic alternatives. These alternatives ideal in the sense of being subjunctively best worlds. Thus if $O\alpha$ is true of an object, it means that it would be better if α were true. It does not mean that it is obligatory for anyone to do something which might make α true.

The dyadic O operator is similar. $O(\alpha \mid \beta)$ is true of an object if the condition given by β determines some non-empty class of deontic alternatives in which α is true. It means that given that β , it would be better if α were true. It does not follow from this that given that β it is obligatory for anyone to do something which would bring about α .

Orthodox deontic logics do not deal with the circumstance that the only way to make α true involves forbidden actions, that making α true is impossible, or that although bringing α about would certainly be better, it is also not *my* obligation but somebody else's¹⁰.

In the literature this discussion is known as ought-to-be vs. ought-to-do (cf. [87]): they are widely *assumed* to be mere reformulations of the same thing, even though it is well known from practical applications that they are not. The assumption that if it ought to be that α (given β) is equivalent to saying that it would be better if α were true (given β) is built into all orthodox deontic logics.

Deontic logics don't represent violations directly. *Legal* obligations do not express the goodness of something but the given that the decision to perform or not to perform the action referred to as obligatory has consequences (being violation or no violation, and the reaction to violation). Interpretation of the violation of the obligation cannot be separated from a theory of decision making and action that answers questions such as what was within the power of the decision maker, what did the decision maker intend, what did he foresee, and what did he try to do? Deontic concepts should be explained in terms of betterness, but also of, firstly, habit formation, and of choice, of beliefs, of decisions, of violations, of sanctions, etc.

Since deontic logics appear to capture the same properties of betterness as preference logics do, but without reference to the social mechanism of norms, there have also been attempts to relate the two concepts (cf. for instance [138, 268, 262, 43, 44]). *Preference logic* (cf. [267, 100, 102, 101]) captures the use of preference (here strongly \prec , and weakly \preceq) and indifference statements (here $=$) between propositions. Against preference logic the same criticism applies: preferences should be explained in terms of choices, intention, beliefs, action, and not just in terms of other preferences.

This connection to decision making and action will be discussed in section 4.7.

The notion of *subjunctive betterness* (as used by [89]) is however a very important aspect of normativity. Logical theories of subjunctive betterness can be investigated on their own, without any further presumption on how norms work, and a number of theoretical issues in AI & Law can be addressed by considering the norm as a statement of subjunctive betterness alone. Both deontic and preference logics take the same stance towards the concepts they formalize, i.e. both express the axiological concept of subjunctive betterness, and both are incomplete formalizations of the thing they claim to formalize. *Subjunctive* refers to the grammatical mood of the type of expressions covered by expressions of obligation and preference. The subjunctive mood is used for subjective expressions, indicating wish, command, possibility, counterfactuals, etc.

In the case of deontic logics this means that an obligation “given β it ought to be that α ” is considered a paraphrase of “given β , α would be better than $\neg\alpha$ ”.

Certain features are characteristic of many deontic logics.

Let $|\phi|$ be the set of worlds w such that $M, w \models \phi$. Central to deontology is the notion of choice:

Proposition 13. Deontic choice $O(\alpha \mid \beta)$: if an agent has the choice between a set of alternatives $|\alpha \sqcap \beta|$ and $|\neg\alpha \sqcap \beta|$ then the agent *should* choose an alternative from

¹⁰This clearly separates legal obligation from ends-oriented accounts of morality.

$| \alpha \sqcap \beta |$.

Clearly one set is better than the other: $| \alpha \sqcap \beta | \succ | \neg \alpha \sqcap \beta |$ if we interpret preference as a relation between sets of worlds described by a proposition, here an OWL Class as defined in section 3.4. This is however not possible in OWL DL: in chapter 6 I will give an account that models axioms constraining betterness assertions $\succeq (i_1, i_2)$ about pairs of individuals i_1, i_2 , that indirectly also covers $\succ (i_1, i_2)$.

In model-theoretic accounts of deontic and preference logics *worlds* are the traditional objects. In OWL DL we are dealing with individuals, but the account is also based on Kripke semantics.

Given the interdefinability of obligation and prohibition, the interpretation of prohibition in terms of sets of worlds follows just like obligation from the principle of deontic choice:

$$\begin{aligned} O(\alpha|\beta) &: | \beta \sqcap \alpha | \succ | \beta \sqcap \neg \alpha | \\ F(\alpha|\beta) &: | \beta \sqcap \neg \alpha | \succ | \beta \sqcap \alpha | \end{aligned}$$

One of the attractive features of the representation in the form of betterness between sets of worlds is that it produces triangles between a context of applicability β , and good ($| \alpha \sqcap \beta |$) and bad ($| \neg \alpha \sqcap \beta |$) alternatives that form a complete partition of the context of applicability. It naturally fits in a graphical representation of taxonomies, and knowledge acquisition methods like the *repertory grid* that combine development of a taxonomy and sorting along evaluative dimensions (cf. for instance [113, 110]).

Beware of interpreting specifically α as an action, and β as a situation: the alternatives may concern both descriptions of actions and situations, as long as situations can be conceived of as productive characterizations in the sense that social and legal norms only speak about situations controlled by human action.

The following are desirable characteristics of a deontic logic of obligations and permissions *as coherent normative positions* (generally based on [195, 268]):

Proposition 14. What is obligatory is permitted: $O(\alpha | \beta) \rightarrow P(\alpha | \beta)$

Proposition 15. The impossible and the meaningless are not obligatory: $\neg O(\alpha | \alpha)$ and $\neg O(\neg \alpha | \alpha)$ are axioms.

Proposition 16. There are no conflicting obligations. The obligations $O(\alpha | \beta)$ and $O(\neg \alpha | \beta)$ are inconsistent: $\neg(O(\alpha | \beta) \wedge O(\neg \alpha | \beta))$ is an axiom. Idem for $O(\alpha | \beta)$ and $P(\neg \alpha | \beta)$.

Proposition 17. . The ordering $| \neg \alpha \sqcap \beta | \prec | \neg \alpha \sqcap \neg \beta | \prec | \alpha \sqcap \neg \beta | \prec | \alpha \sqcap \beta |$ satisfies the propositions $O(\alpha | \top)$, $O(\beta | \alpha)$, $O(\neg \beta | \neg \alpha)$.

Since law *will* create conflicting obligations, a proper deontic logic which models normative positions is not necessarily also an appropriate tool for representing legal rules. For a deontic logic for legal purposes there are additional design requirements not addressed in detail here. Most importantly we have to deal with conflicting obligations.

In chapter 6 the representation of permission is derived from the following interpretation:

$$P(\alpha|\beta) : |\beta \sqcap \alpha| \succeq |\beta \sqcap \neg\alpha|$$

The translation of a permission to a statement of weak preference instead of indifference ($|\beta \sqcap \alpha| = |\beta \sqcap \neg\alpha|$), as has been argued by for instance [212, 239], is a point of contention. Indifference is incompatible with the intuitions of deontic logicians (cf. [2]); If something is obliged, then it should also be allowed. The asymmetric \succeq statement leaves room for a prohibition or obligation (\prec) and retains the information that the represented permission explicitly allowed a case and not its opposite, on which it was silent.

It is perhaps extended exposure to alethic modal logics and law that causes the intuitions. There is empirical evidence that children generally attribute an attitude of indifference to others who express a permission (cf. [170]).

A norm system that only contains $P(\alpha|\beta)$ is not a well-formed norm system at all since it serves no purpose in guiding and evaluating behaviour. Operator P serves no real purpose in a deontic reasoning system that does not allow for conflicts between norms in guiding and evaluating behaviour. The operator P cannot be understood in any other way than a superfluous utterance stating indifference towards $|\beta \sqcap \alpha|$ and $|\beta \sqcap \neg\alpha|$ if it is evaluated without other normative expressions as context. The P only becomes relevant if:

- it conflicts with an obligation or prohibition existing in the context of discourse, and
- it is used to cancel the conflicting obligation or prohibition.

The explicitly stated permission clearly has another function than the dispositional permission inferred from the absence of knowledge of a prohibition (the *silent* of Valente's interpretation). Since permissions are usually uttered with the explicit intention of amending a specific obligation, and are often found nearby in the same legislative text, it makes sense to 'localize' them to some extent by adding the asymmetry. The asymmetry is also necessary for using the representation in combination with common deontic reasoning systems.

Some broad permissions can perhaps be interpreted as symmetrical. It is for instance generally accepted that freedom of expression i.a. includes a strong permission to keep your opinion to yourself. We might add a fourth dyadic operator for freedom of choice (liberty) with some limited applications in the representation of legislation:

$$L(\alpha|\beta) : |\beta \sqcap \alpha| = |\beta \sqcap \neg\alpha|$$

Obviously the same effect can be reached by asserting two permissions.

Putting the axiological interpretation in terms of betterness, and the institutional interpretation in terms of constitutiveness of legal facts (Valente's allowed or disallowed) together results in the structure described in figure 4.2. This is the target for representation of normative rules in section 6.2, after we have laid a foundation for the representation of legal rules in terms of institutional and constitutive rules in chapter 5.

A legal norm is an a priori obligation or prohibition (i.e. not a positional one), it applies to a certain case, allows a certain case – the allowed case – and disallows a

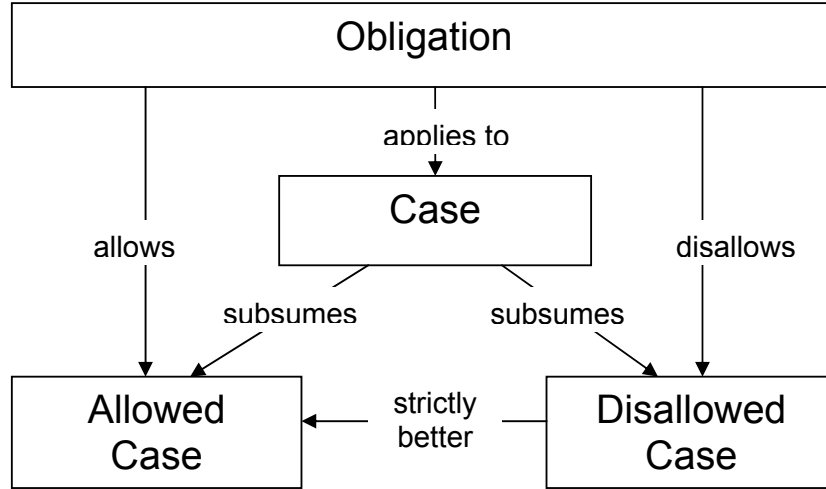


Figure 4.2. An entity-relationship diagram describing the salient structure of obligations and prohibitions.

certain case – the disallowed case. Instead of saying that the norm disallows the case, we can also say that the case violates the norm.

The allowed and disallowed case are both subsumed by the case to which the norm applies. Besides that they by definition form a complete partition of the case to which the norm applies, i.e. all cases to which the norm applies are either allowed cases or disallowed cases. This is true of the obligation and the prohibition: they are simply two different ways to put the same thing into words: the obligation explicitly mentions the allowed case, and leaves the disallowed one to common sense inference, while the prohibition makes the disallowed case explicit.

The nature of the betterness relation between allowed and disallowed case is from the perspective of the addressee justifiable from the consequences of violation of the legal norm, while the legal norm is itself an expression of a preference of the legislator.

In terms of constitutiveness and betterness norm $n = O(\alpha \mid \beta)$ can be interpreted syntactically on the following basis in OWL, as will be shown in section 6.2 later in this book:

1. the three relevant *cases* correspond with OWL classes, β , $\alpha \sqcap \beta$, and $\alpha \sqcap \neg\beta$;
2. normative rule n is an individual, which disallows certain individuals, modeled by OWL assertions $\text{disallows}(n, i)$, etc.;
3. the normative rule is however also a nominal concept $N \equiv \{n\}$ of which we can assert terminological OWL axioms such as that of all things disallowed by n the proposition $\beta \sqcap \neg\alpha$ is true: $\{n\} \sqsubseteq \forall \text{disallows}.\beta \sqcap \neg\alpha$;
4. normative rule n orders certain individuals i_1, i_2 by betterness, modeled by OWL assertions $\succ_n(i_1, i_2)$, etc.; and
5. we can assert terminological OWL axioms about this betterness interpretation

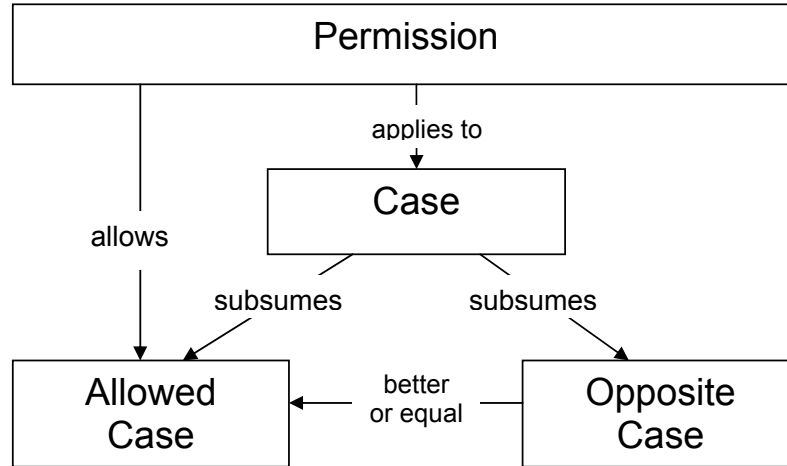


Figure 4.3. An entity-relationship diagram describing the salient structure of a permission. Obligation and prohibition are subsumed by permission.

of n , such as that in situations in which the proposition $\alpha \sqcap \beta$ is true all better or equal situations are situations in which $\alpha \sqcup \neg\beta$ is true, or $\alpha \sqcap \beta \sqsubseteq \forall \preceq_n \alpha \sqcup \neg\beta$.

The permission, shown in figure 4.3, is only slightly different. The permission allows something, but it doesn’t disallow anything. The logical complement of the mandated case is here simply called the opposite, following existing practice by some authors (cf. [266, 156].

A desideratum for the OWL representation in section 6.2 is that the obligation is subsumed by the permission. In addition, one can easily imagine a more generic subsuming concept that just “applies to a case” for constitutive rules in general, and we will indeed develop such a notion in section 5.2.2.

The analysis of normative rules cannot, however, stop here. Obviously, the function of normative rules cannot be considered without appealing to the notion of *agency*. The normative rule does not merely express relative goodness of certain states of affairs, but also directs agents (not) to bring them about.

The betterness interpretation as developed here orders worlds, or in OWL terms individuals. The question is however what concept describes the domains of the \prec relation. Is it a situation, an action, or a plan?

4.7. Agents and Action

Agency is not a subject that should be considered part of AI & Law: most accounts of action and planning are written outside the field of AI & Law, and are often designed with automated planners – that construct a plan to achieve a certain goal state from some initial state for a domain with certain properties – in mind (cf. generally [25,

29, 107, 58, 85, 4, 178, 3, 137, 213, 261]). From a Semantic Web perspective it makes sense to combine a generic mechanism for deontic reasoning with a theory of action, events, or plans designed for some other purpose.

These theories can vary considerably in how they deal with time, with change, etc. Still, we cannot satisfy ourselves with an account of deontic reasoning that abstracts the thing being qualified to some opaque proposition. Doing so gives rise to odd analyses of what norms mean for human action.

We have to make at least some assumptions about agency and change, and relate these to normative positions like being in a state of obligation to do something. An important concept in this respect is the distinction in legal theory between *ought-to-do* and *ought-to-be* representation, discussed in the next section.

In a nutshell, this chapter tries to account for two major uses of legal rules:

Planning to perform a task one sets oneself; to bring about beneficial legal facts, while avoiding detrimental ones; and

Situation and action recognition to infer one's own legal position, and what legal facts others bring about, are going to bring about, and unsuccessfully attempted to bring about.

Planning involves generating and comparing alternative plans, while situation and action recognition involve generating and comparing explanations. The planning perspective also prominently involves time and change.

The purpose of this section is to point out some constraints that the conceptualization of intelligent behaviour implicit in the rules themselves impose on the ways in which we can conceptualize agency and change. It is not intended to set a golden standard for the conceptualization of planning and change itself.

4.7.1. To Be or To Do

A central subject in the representation of obligations is the *ought-to-be* versus *ought-to-do* debate (cf. [74]): does the obligation prohibit a state of affairs or the bringing about of a certain type of change? What role does intention play in this? If obligation cannot be reduced to mere subjunctive betterness, as for instance [74, 89] argue, we have to account for what an obligation tells us to do.

Important for knowledge engineering is a derived issue: is there a logical relationship between these two possible formulations that allows us to translate between them?

The key question for this chapter is however: to what kind of objects are the constitutiveness and betterness interpretations applied? Situations, actions, plans, or something else?

[265, 184, 152] argue for a distinction between the following subproblems of enforcing norms:

1. the norm proper, which simply derives *{allowed, disallowed, silent}* from a situation description,
2. the causal connection between the conduct of a person and the qualified situation,

3. the fault implied by the conduct of that person, and
4. the attribution of responsibility, usually but not always to that person¹¹.

The discussion in [152] gives the best overview of the mechanisms involved, and stresses that causation itself is the outcome of ampliative inference explaining the occurrence of an event by a process explained by an earlier or simultaneous event. This causal model is an abstraction of underlying processes. Agent causation simply extends this to mental processes.

Since the end result is an explanation in terms of *agent causation*, or the lack of it, and the legislation itself can be understood in terms of agents “bringing about” i.e. causing certain occurrences, this mechanism of explanation is of minor concern here: legislation typically addresses the explanation, and not the underlying processes.

There are two markedly different “styles” of legislative drafting that are clearly both encountered in legislation: the productive style, which focuses on unwanted results, and the behavioural style, which focuses on ease of recognition of unwanted behaviour. Note that the productive style does obviously attribute it to behaviour: the ought-to-be style describes the goodness of the product of our actions. It deals with states of affairs generally caused by humans, but omits specification of their *role* in it. The ought-to-do style directly describes the action. As pointed out in section 4.1, action recognition presumes recognition of the scripted role the agent assumes and ascription of intent.

Generally speaking, in settings involving interaction between a human and a machine, or a design of one, it is easier to give a productive characterization of what the legislator doesn’t want: e.g. *the carrying capacity in kg of the elevator should be at least 100 kg times the number of people you can reasonably fit into the volume of the elevator*. It is of course possible to give a behavioural characterization of the same constraints, but this approach can be clumsy and unintuitive.

In these settings we are usually dealing with what is often called a *classical environment* that is fully observable, deterministic, with a finite number of discrete alternatives at any point, and where the system or design being manipulated is only changed by the planning agent. If the system or design violates a rule, it is its owner, operator, or designer that did it.

In environments where one or more of these features are not present there is a less predictable interaction with the environment. In this kind of setting we have failed attempts, unforeseen products of acts, and ambiguity about who caused something. Because the intention of the law is to a large extent to influence people’s behaviour, it is more practical from a compliance and enforcement perspective to give behavioural characterizations of behaviour and to point directly at the seller, the driver, etc. Behavioural characterizations are generally speaking easier for the addressee to learn and adopt, and violations of norms with behavioural characterizations are often easier to recognize and prosecute.

Because the abstract goals of the legislator often have a productive nature (for instance reducing the number of people that die in traffic, or reducing conflicts over the possession of goods, or raising money for some purpose), behavioural norms often

¹¹For instance liability of a parent for the conduct of a child.

feel like approximations of what the legislator really wants of the addressees of the legislation. Behavioural characterization increases precision in enforcement and compliance, by making the things to which it applies easier to foresee and recognize, but at the expense of precision in describing the policy goals of the legislator.

Productive characterization creates problems for the addressee if the addressee does not foresee the consequences of his actions reliably in an environment that is only partially observable and stochastic. This is for instance often the case if other human beings are involved. If the addressee perceives a lack of control over the situation, enforcement of legal norms has little effect.

The full mechanism in [152] appears to be conceptualized *explicitly* in the sources of law only in cases of grave importance, for instance responsibility for someone else's death. Here the legislator not only addresses unintentionally causing death, but also failed attempts at causing it.

The fundamental difference between an ought-to-do characterization and an ought-to-be characterization is in the omission in the ought-to-be norm of an action description and an agent identified by role (e.g. the driver coming from the right, the thief who is taking a good that wholly or partially belongs to another). The problem with the ought-to-be characterization is therefore mainly in identifying the agent who *caused* the state of affairs or failed to prevent it from occurring, while in ought-to-do characterizations this is merely a matter of action recognition.

While it is the state of affairs that is considered to be disallowed (following [152]), it is only an agent who will be, for his actions or his failure to act, held responsible for its occurrence. For the agent this means that when he foresees that his action may end up in an prohibited state of affairs, the problem is whether responsibility for the state of affairs will be attributed to him or to someone else.

This is not trivial. It is not necessarily the case that the one held responsible for an undesirable situation should also be the proximate cause of the situation in which the violation arises. You can't for instance simply reduce ought-to-do to ought-to-be. An example: Imagine someone offers exclusively to you a painting for sale, with the condition that the offer is valid until next Tuesday. You have a claim-right to buy the painting, in the sense that the counterparty has a duty, towards you, to sell you the painting *if* you accept the offer before next Tuesday: the behaviour prohibited consists of not selling it to you while you have accepted the offer. The counterparty then sells the painting to someone else, before Tuesday. After he sold it, you accept the sale. You are now in the situation prohibited by the obligation, and your conduct is the "proximate" cause of the situation by expressing your intention to buy the painting, but not the legally relevant cause. You are not responsible, because the duty wasn't yours but the seller's: the fault is with the seller.

In this case, for you, the obligation does *not* mean anything in terms of the betterness interpretation. While it is possible to claim that it is the occurrence of the situation that is allowed, its interpretation in terms of obligation cannot be based on situations.

Proposition 18. The mere fact that a *situation* appears prohibited does not necessarily mean that a plan that brings it about is a bad choice for a specific agent. It depends on whether you will be held responsible for it. If however the agent conceives

of the plan as an alternative in a planning problem, and the plan is based on the assumption that other agents will behave *normally*, i.e. following the prevailing norms, then it is very likely that choosing to execute that plan is a bad choice.

Even if we consider the attribution of responsibility to agents as unproblematic, and subscribe to a simple deterministic model of states and transitions, the relation from behaviour to product is not always immediately obvious. d’Altan et al. (viz. [87]) for instance offer the following hypotheses for the relation between ought-to-do characterizations and ought-to-be characterizations:

1. A state is obligatory if and only if it is the result of an obligatory action.
2. A state is obligatory if and only if all the actions that lead to the state are obligatory.
3. A state is obligatory if and only if it is forbidden to undo it.
4. A state is obligatory if and only if all the actions that are necessary¹² to bring about the state are obligatory.

None of these alternatives is without counter examples, as [87] point out. It is certainly not obvious that all ought-to-do norms can be translated to ought-to-be ones; The other way around is more plausible.

There is another argument in favour of actions as the subject of obligations.

In addition to behavioural and productive characterization, the *attempt* forms a third category: characterization of display of undesirable intent in itself, regardless of the form in which it is executed successfully. This category is problematic, and therefore usually avoided if possible. There is a strong taboo on “ruling over the consciences” of citizens, and a practical reason to avoid it is obviously that it requires mind reading on the part of the court.

Still legislation with a strong moral tone, in particular criminal law, does involve a certain degree of “mind reading” when dealing with things like attempts (to kill someone), premeditation, or the distinction between intentional and unintentional killing. Because the court cannot actually read minds, it has to infer apparent intention from actions. To do this, it has to judge what the agent foresaw as the consequences of his actions, and this requires ascribing knowledge to agents. This isn’t really helpful, as it just moves the mind reading problem to another aspect of the operation of minds.

The solution is unsurprisingly found in the ambivalent nature of norms: it is possible to set standards for what reasonable people *should* know. There is also a related legal concept: the knowledge of the *man on the Clapham Omnibus* (cf. [203]).

How to realize this in KBS is hardly a problem: it is usually taken for granted that the dispositional belief set of some agent (cf. section 3.3) is the logical closure of its individual beliefs with some universal body of shared knowledge. The idea that we share a body of knowledge is an assumption behind ontology modeling. It is indeed a necessity that people who understand each other share a body of knowledge. In addition to shared ontological knowledge, people must also be assumed to share knowledge about what normally happens, as alluded to in section 4.5.2.

¹²I.e. performance of the actions necessarily leads to the state, and non-performance of the actions necessarily prevents the state.

The attempt is an unsuccessful execution of a task. The attempt is therefore an action according to the principles decided in in section 4.3, while the situation that results from it can be left completely unspecified.

One can however also apparently be held responsible for one's non-action. Of this scenario we know two variants:

1. One has an obligation to act and doesn't during some period; and
2. something that is disallowed happened because one did not act to prevent it (*negative causation*, cf. [184, 152]).

The first case is a standard feature of any combination one makes of a deontic logic and causation (addressed in 4.7.2).

The second case only happens if one is presumed to have been aware of the possible undesirable occurrence and didn't attempt to prevent it, i.e. demonstrated a lack of intention to prevent the occurrence through one's choices. Assuming that agents are always executing some task they set themselves, the undesirable occurrence was "on the radar" during the planning of that task and therefore part of the conceptualization of the situation for which one made a plan. Letting the undesirable occurrence happen is therefore part of the task specification and therefore intentional: here too one is in the end responsible for what one did do.

The *action* is what links undesirable changes *to* some situation, undesirable changes *from* some situation, undesirable continuation of a situation, and unsuccessful attempts directed towards these. The subject of the norm is therefore naturally the *action*.

Proposition 19. The subject of a normative rule is an *action*.

4.7.2. Agent Causation and Time

A familiar framework for explaining how stuff is brought about is *event calculus* (cf. section 3.5.3). By acting, an agent either initiates or terminates certain *fluents*, for instance a:

1. *state* of an object; for instance the state of being married, or the state of obligation of the counterparty to deliver certain goods,
2. *situation*, i.e. a configuration of objects; for instance two people being married to each other, or
3. perdurant *object*; for instance a marriage, or a private company limited by guarantee.

Sartor in [244] (table 2) lists a number of *normative conditional* patterns based on the distinction between initiation (action initiates fluent), termination (action terminates fluent), and *emergence* (fluent entails fluent) where the consequent is always a legal fact. These facts are further classified into normative positions (e.g. my obligation to terminate the traffic violation initiated by the traffic participant overtaking me), qualifications (e.g. being a married person), and the existence of legal things constituted by something else (e.g. the marriage that exists between me and my wife).

All of these are obviously constitutive rules. The difference between qualification (married) and existence (a marriage) in this sense is primarily a difference of descriptive

vocabulary and not of substance: reification of relations in a knowledge representation and nominalization of verbs or adjectives in natural language are motivated by technical and presentation considerations, not ontological ones.

Initiation and termination are generally encountered in real life as pairs – a *change* operating on the same substrate, which can be an object or some configuration of objects (a situation, scenario, plan, system, etc.):

$$\begin{aligned} \text{CHANGE} &\sqsubseteq \exists \text{initiates.STATE} \sqcap \exists \text{terminates.STATE} \\ &\sqcap \exists \text{substrate.OWL:THING} \sqcap \exists \text{extension.PERIOD} \\ \text{STATE} &\sqsubseteq \exists \text{substrate.OWL:THING} \sqcap \exists \text{extension.PERIOD} \end{aligned}$$

The change terminates a state and initiates a state of the same thing. State (for “stative”) is taken to be the generic type subsuming the previous categorization of ways in which fluents are described.

Observe that the combination of the deontic operator ‘it is obligatory that’ (O) from section 4.6 with a simple intentional action operator ‘agent i brings about’ (E_i) would give rise to four atomic types of obligations (cf. generally [250] for a discussion):

1. it is obligatory that i brings about C ;
2. it is obligatory that i brings about not C ;
3. it is obligatory that i does not bring about C ; and
4. it is obligatory that i does not bring about not C .

As noted earlier, action presumes the recognition of plan execution by i . Let’s for the moment set aside the problem of how i is going to (know whether i can) bring about C , and how one determines whether i did, didn’t, attempted it, etc. The difference between ‘ i brings about not C ’ and ‘ i does not bring about C ’ requires some further analysis.

Obviously bringing about not C is only meaningful if C is the case: we can reformulate this to *terminating* C (whenever it may arise), while not bringing about C is about *not initiating* C , and C is a state, some occurrence that occupies a time interval or *period* on the canvas of time, that is initiated by some change and terminated by another change. Similarly, we can talk about *not terminating* C (whenever it may arise), and about *initiating* C .

The agent can be held responsible for the fluents he initiates and terminates: this is the easy case. He is also responsible for the fluents he was able to initiate or terminate but didn’t. In the Netherlands it is for instance permitted for a pair of bicyclists to ride next to each other. If a third traffic participant overtakes them, this traffic participant initiates a traffic violation, against the obligation to keep to the right. At the same time jurisprudence shows that each one of the three traffic participants can be held responsible for the traffic violation, since each of them is able to terminate the traffic violation. Current jurisprudence apparently interprets the obligation as pertaining *both* to initiation and to termination, but this does not follow from the phrasing of the involved obligation and permission which simply addresses the stative pair of bicyclists riding next to each other (in other words, an “ought-to-be” characterization).

This should be kept in mind when interpreting ought-to-be descriptions: does it pertain to initiation, to termination, or to both.

Normative rules are usually about bringing some change about. There is a general pattern explaining how legislators will typically phrase normative rules:

O : some change \succ the continuation of the sketched situation;

F : continuation of the sketched situation \succ some change; and (less markedly)

P : some change \succeq continuation of the sketched situation.

Norms are only concerned with initiating and terminating C , and do not have any bearing on any actions that neither initiate nor terminate C . None of these formulations tells me anything about bringing about unrelated fluent D acting on another substrate.

Only in the case of directives to *not* initiate or terminate C we can determine directly whether a possible violation has taken place: it takes place if we *do* initiate/terminate C , i.e. when a change occurs. Also if the obligation specifies that C must be terminated *before* D is initiated, violation takes place if D is initiated while C is still the case. Directives to initiate or terminate C are however assessed against *deadlines* (cf. generally [97]), and only the passing of the deadline gives rise to violation. The obligation to initiate C may for instance well be irreconcilable with a course of action that consists of first initiating D and only then C , but this does not have anything to do with the position of the initiation of D relative to the initiation of C in time.

The implicit deadline leads to a bit of a problem if we consider this in the context of things allowed and disallowed by a norm. Take as an example the obligation that i terminates C ; This means the following in terms of normative qualifications:

1. i terminating C is allowed;
2. the continuation of C by i (beyond some unspecified deadline) is disallowed.

The representation of case 2 is obviously a bit problematic, because it addresses the non-action of ‘doing nothing about C ’. Exact identification of the intended logical complement of terminating C is hard.

Let it for instance be prohibited to drive with a vehicle without functioning headlights between dusk and dawn. In practical terms this is supposed to mean that 1) one has to check whether one’s lights function before one starts driving, and 2) that one has to switch the lights on before dusk. One may discover a headlight is broken while on the road after dusk sets in¹³. One is now in the state which is disallowed, without consciously having decided to initiate it, and one is obliged to terminate it.

Leaving the road to repair one’s headlights is usually going to involve some more driving on the road, to find a suitable place, to buy the light bulb at a gas station, or to go to a garage. Since no deadline conditions are given, intuitively, only action betraying the *intention* to continue the violation would *clearly* constitute a violation. If one gets pulled over by the police one can get fined, but it is for instance unclear *how many times* one can get fined for substantially the same fact: the fact that one is fined for not terminating C does not mean that the obligation disappears. It would however be unreasonable and unfair for the police patrol to stay around for another

¹³Since lights tend to break when switched on, the earlier check doesn’t help.

opportunity to give a fine and they generally don't. It is also reasonable to not give a fine if the driver can repair the light directly. In other scenarios other factors determine enforcement decisions: enforcement decisions depend on enforcement policy and, in the absence of such policy, private considerations.

Proposition 20. It is not possible to determine whether an obligation to bring about a change has been violated, unless one sets a deadline. Obligations to bring about a change *before* some other change present no problems. Reasonable enforcement policies are based on ascription of intention to behaviour while the obligation exists.

In a logic of agency which treats actions as instantaneous changes of state occurring at discrete time steps (i.e. Hoare logic, dynamic modal logic), it is natural to consider the next state as the implicit deadline. Combined with the presumption that, given full knowledge of the initial state and a known menu of possible actions, the state following execution of an action should be predictable with absolute certainty, we can axiomatize *obligation* completely in terms of agency, like the proposals in [87]. We should keep in mind, however, that the verity of these axioms depends on the conceptualization of the planning domain, and not on the properties of the concept *obligation*: it doesn't shed any light on the nature of obligation.

Representing a normative rule involves some common sense judgment on the part of the knowledge engineer. How complicated a representation needs to be depends i.a. on the following:

1. Does the normative rule demand that one displays the intent to comply, i.e. does it take into account the possibility of failure of attempts?
2. Does the rule require continuation or change?
3. If it describes a change, is it phrased as an initiation or a termination?
4. If it requires a change, is there an indication of a deadline or a *before* condition?

We can identify a number of ingredients that should be covered, although we cannot solve the deadline problem. For the representation of normative rules, apply the following principles:

1. Normative rules apply to action;
2. Action causes changes;
3. Action, whether conceived of from the perspective of the actor or recognized by a bystander, is situated, i.e. takes place in a situation; and
4. A situation consists of a limited number of participants, it is not a (state of the) world in the logical sense (cf. section 3.4), but a conceptualization of the context in which the action takes place.

The direct connection of *situation* to action is inspired by [261, 15], and the representation of discourse context in general in computational linguistics (for instance [227]). It deviates significantly from AI planning literature, which is largely based on the assumption that 1) world and situation are the same thing, and 2) that situation + action results in a new "situation". This works well for planning, but is not terribly realistic.

Instead an agent reconceptualizes the situation in the context of planning the next action: since this is done in the knowledge of one's previous action, and one may be

acting still in the same role, it is not surprising if the new conceptualization borrows participants from the old one. Since we have no particular application in mind, it is wise to take the conceptualization in the source of law at face value.

Actions are situated, performed by an actor, and cause changes:

$$\text{ACTION} \sqsubseteq \\ \exists \text{situation.OCCURRENCE} \sqcap \exists \text{causes.CHANGE} \sqcap \exists \text{actor.AGENT}$$

As always, this is the simplest possible conceptualization, ignoring issues like decomposition of both actions and situations. The obligation to repair a broken headlight – without a clear enforcement policy – covers the following three concepts:

$$\exists \text{situation.BROKENHEADLIGHT} \quad (4.1)$$

$$\exists \text{situation.BROKENHEADLIGHT} \sqcap \forall \text{causes.}\neg \text{HEADLIGHTREPAIRED} \quad (4.2)$$

$$\exists \text{situation.BROKENHEADLIGHT} \sqcap \exists \text{causes.HEADLIGHTREPAIRED} \quad (4.3)$$

The obligation applies to 1, disallows 2, and allows 3. Disallowed is any action performed in a situation which involves a broken headlight and does not involve fixing it. To find an acceptable resolution to deadline issues we would have to address decomposition of actions, for instance:

1. Disallowed is any action performed in a situation which involves a broken headlight and does not involve fixing it, and is not part of an action that involves fixing it.
2. Disallowed is any action performed in a situation which involves a broken headlight and does not involve fixing it, and is not an execution of a task that involves fixing it, or of a subtask of a task that involves fixing it.

These kinds of refinement are either based in common sense or in jurisprudence. The scenario of negative causation, although superficially similar, is different. In this case a *future* change is implied by one's conceptualization of the situation; An action likely to be deemed wrong is for instance:

$$\exists \text{situation.IMPENDINGDISASTER} \sqcap \forall \text{causes.}\neg \text{PREVENTDISASTER}$$

It may seem strange to include expectations about the future as part of the situation one acts in, certainly if one conceives of situations as is usually done in AI planning, but it is no stranger than including expectations about the past: if we see someone coming towards us with a blood-covered chainsaw, our explanation of this event will cover both the past (to explain where the blood comes from) and the immediate future¹⁴. The notion of normative position, which can also be conceived of as being part of the situation in which one acts, is also often essentially about *future* possible consequences of one's actions.

¹⁴The explanatory power of a scientific theory of the past, like evolution theory, is for instance in its prediction of future observations about the past, i.e. missing links yet to be found.

4.7.3. Positions and Power

Interesting are the positional uses of obligation, called *normative positions*. The most obvious of these is the state VIOLATION initiated by an action that is disallowed:

$$\exists \text{disallowedBy.NORM} \sqsubseteq \exists \text{causes}.\exists \text{initiates.VIOLATION}$$

Other familiar ones are systematically related to corresponding normative rules, but are secondary to the reading of normative rules as being constitutive of the qualifications $\{\text{allowed}, \text{disallowed}\}$. Given the existence of a norm “given α, β it ought to be that γ is initiated” and me being in a context that entails α , I have a contextualized obligation to initiate γ if β .

The relation between this type of concept and the original norm is similar to “being in a position in which I have to get ingredient a ” when I am preparing a dish that requires ingredients a, b, c and I have only b and c . If I believe I cannot obtain ingredient a , my position would be “not being able to make the dish”. This is a planning concept (and a fluent that exists only in my mind, or that of others) that has no direct bearing on the recipe.

The same applies to norms: the contextualized obligation is not the original norm, but a state relative to the norm that only acquires special meaning by adding a theory of what I am able to bring about and what I am trying to do. It is possible but not very useful to automatically derive positional obligations. In scripts, as planning concepts, they have their uses as a shorthand for a compound action, for instance an action in a situation α that brings about the initiation of β but not of γ is not allowed by n :

$$\begin{aligned} \exists \text{situation}.\alpha \sqcap \exists \text{causes}.\exists \text{initiates}.\beta \sqcap \forall \text{causes}.\forall \text{initiates}.\neg \gamma \\ \sqsubseteq \exists \text{disallowedBy}.\{n\} \end{aligned}$$

While normative rules apply to action, positional concepts also often follow the state constitutes state pattern, for instance:

$$\text{FIREALARMSIRENWAIRS} \sqsubseteq \exists \text{constitutes.FIREALARM}$$

Positional concepts play a role mainly as a description of the result of the application of a constitutive rule. They also occur as conditions to constitutive and institutional rules. A special type of condition is the (legal) power or competence. In common law jurisdictions the term power appears to be preferred (cf. Bentham, Hohfeld, and Hart), while civil law jurisdictions commonly call it competence (cf. [69]).

The *power* to bring certain things about refers to a certain legal qualification in the constituting base that applies to the agent that gives the ability to bring about a legal fact. “Only parliament declares war” for instance means that being the parliament is a necessary condition for the initiation of the state of war. It is distinguished from other similar potentialities (“Only people who have money can buy stuff”) by the fact that the qualification is a fact within the institution.

The agent is able, competent, or it is within his power, or he has the potential, to bring certain things about. Almost every agent is for instance able to open and close an unlocked door, but most agents are not able to open a locked one. This is *ability*, and was already indirectly addressed in section 4.5.2: to have the ability to do something in a certain situation is to meet the conditions for filling an agent role. Some agents are able to marry a pair of other agents, but most are not. Since marriage is an institutional status (nowadays), this is (legal) *power* or *competence*.

The abilities and powers of a person are limited to the agent roles for which he meets the necessary conditions *in the situation*. Agent roles are *stative*, and in that sense part of the situation in which an action is performed: it is however common to distinguish states that specifically pertain to the agent (e.g. police officer) from those that do not (e.g. dark), which are rather conceived of as conditions for executing the task, and relatively shortlived ones (e.g. unarmed) from longlived ones (e.g. police officer).

Some authors have tried to explain power-conferring rules as a variety of normative rules (notably Von Wright, Bentham, Cornides, Kanger, Lindahl, and Kelsen), while others explain them, as I do here, in terms of constitutive rules constraining institutional action (Hart, Ross, Searle, and Bulygin in [69], which is also the source used here for the classification of authors by the position they take on this subject).

As pointed out in 4.4, we usually speak of a power if the legislator intended to create a way for a certain class of agents to bring about the legal fact. Attribution of power is captured by standard constitutive rules, if these have legal conditions that pertain to the agent role, i.e. for instance:

$$\text{MARRY} \sqsubseteq \exists \text{actor}.\text{HASPOWERToMARRY} \sqcap \exists \text{causes}.\exists \text{initiates}.\text{MARRIED}$$

The marry action is performed by an agent with the power to marry and causes the initiation of the state married.

4.8. Interaction Between Agents

The mere publication of a formal representation of set of constitutive rules does not make a functioning legal system. Implicit in the act of legislating is the implied threat to *react* if violations take place. The *Functional Ontology of Law* in [265] therefore posits a special category of legal knowledge that covers this function: *reactive knowledge*. The institution however does not depend on an ability to “punish” itself: it reacts to violations by either obliging, permitting, or empowering others to react. Essential to the functioning of the institution is that there is a critical mass of agents who are willing to follow these directives or use these means provided to punish others.

Tort law for instance depends on payment of damages to function as a sanction of the injuring party and an incentive to report the violation for the injured party, and police officers follow directives because they are paid to do so with taxes collected (using the law) i.a. for the benefit of having police officers who react to violation of criminal law directives.

The following is for instance a somewhat simplistic constraint¹⁵ on the power to arrest:

$$\text{ARREST} \sqsubseteq \exists \text{situation. OFFENCE} \sqcap \exists \text{actor. POLICEOFFICER}$$

Arrests can be performed in a situation of offence, and by a police officer.

The institution works if it is possible to arrange normative order in such a way that reacting to one agent's violation of the law is in the interest of another. Very often a role is given to the "victim" of the violation, but it is important to note that no appeal to *restorative justice* (i.e. the notion of norm violation as an act against another individual, who should receive some kind of restitution) is needed to explain why the victim usually takes on this role. Very successful and efficient (but morally unappealing) enforcement systems have for instance been based on the principle that the agent who brings the case (and often also the suspect) in for adjudication receives property of the agent violating the law after he is executed or enslaved. The legislator counts on self-interest and commitment to the public interest.

Reaction of course subsumes not just punishment, but also the mere reporting of the case, gathering evidence, optionally apprehending suspect(s)¹⁶, adjudication, and any other supporting activities. It is customary, for obvious reasons, to separate the role of adjudicator (who authoritatively interprets the law and applies it to the case) from the reaction *per se*.

As pointed out before the whole system depends on nothing but constitutive rules, but these rules are arranged in such a way that the addressed population of agents makes enforcement work by organizing punishment of each other, and collecting the necessary funds for employing professionals to make the system run smoothly.

A simple example of this mechanism can also be seen in international paralegal frameworks like ship *classification societies*: ship owners voluntarily pay a subscription fee to a classification society, which in turn uses the collected funds to draft rules and send surveyors to ships to enforce compliance with these rules. The rarely used sanction consists of suspension or loss of "class", which is a problem for the ship owner because port authorities generally demand classification with a reputable classification society. authoritative interpretation of the rules in final instance is left to some court.

[265] is obviously right in noting that there is a functional distinction between norms and other rules that regulate normal interactions and those that regulate reaction sanctioned by the legal system. Contrary to some of the other distinctions made there (for instance between world knowledge and normative knowledge), reactive rules are however in no way distinguishable from other categories of rules by their content. It is their intended function in decision making which identifies them. Another functional category in [265] to which this observation applies is *metalegal knowledge*, which contains the rules of adjudication.

Proposition 21. The judge, police officer, parliament, etc. just follow the rules of the institution. There is no distinguishing *formal* criterion that sets apart reactive rules from other rules. The distinction is purely functional, and often not easy to make.

¹⁵Offence is simplistic.

¹⁶This step only becomes important if the foreseeable punishment justifies running away.

4.8.1. Transactions and Interests

The previous section strongly suggests that the *interests of others* play a large role in deciding what you can do, and that the law is usually set up in such a way that your violation of an obligation may give rise to another person's permission or power to do something to you which is in their, and contrary to your, interest. An example: your obligation as a driver to yield at intersections for drivers from the right, is correlative to their right of way. This may mean two things:

1. Your violating against the obligation may be a precondition for some action the other driver can take against you. Typically, the violation of a traffic rule would make you for instance liable to paying for the damage if an accident happens. This is a specific position which is associated to him being the victim of your action, and it is in his interests to use this position.
2. More important in daily life, the other driver may proceed with the *legitimate expectation* that you will yield. The other driver will typically prefer driving on over stopping, and will expect you to stop in order to comply with your obligation.

These two options need not be correlated although they often are. A police officer can *react* to the traffic violation by fining you, by virtue of being a police officer, without being the victim of your action. The victim of your action is not in a legal position to react to you if no accident happens.

Most normative positions are two- or more-sided in the sense that there is or are possible victims of your noncompliance. But how do we decide that some bystander is a victim of your noncompliance, or a beneficiary of your compliance?

A number of implicit interests are involved in the example, that can be represented in analogy with the subjunctive betterness interpretation of norms:

1. The other driver usually prefers not stopping over stopping, and therefore you behaving normally over you not behaving normally.
2. Both drivers prefer no accident over an accident.
3. The other driver usually prefers you paying for the damage over you not paying for the damage in the case of an accident.
4. You prefer not paying for damage over paying for damage.

The legislator is making certain assumptions about the preferences of typical bystanders depending on their agent role *agent*. In legislation these usually remain implicit.

Imagine an agent who is solving decision problems all the time in a situation involving other agents who might do unpredictable things based on their own decisions. In his decisions he has to take into account his legal position. In the first place, this means that there are certain things that he cannot efficiently bring about, because they are disallowed. But since these norms also apply to the other agents, there are also certain things he *can* efficiently bring about by taking into account their obligations, and assuming they behave normally. Crossing intersections while somebody else is coming from the left for instance, in the expectation that they will yield, or sending

someone the product they ordered, in the expectation that they will make a payment upon receiving it.

Proposition 22. The law does not only constrain action: It also creates abilities.

The true real life beneficiary of an obligation is the one who has a preference for someone else meeting their obligations. The other driver on the intersection is therefore typically the beneficiary of the obligation to yield. This is to be distinguished from another function of explicit assignment of the role of beneficiary by the legislator: In traffic rules the other driver will not be explicitly recognized as the holder of right correlative to obligation, and in criminal law somebody may be explicitly recognized as beneficiary without actually having an interest in it: think for instance of euthanasia as murder. In many cases this fiction is used to attribute *potestative right* to the beneficiary to take some action following noncompliance, while in other cases – like violent crime – the notion of a victim is purely fictional, and intended only for the justification of the law itself.

Proposition 23. The legal institution ascribes preferences to persons depending on the agent role they adopt.

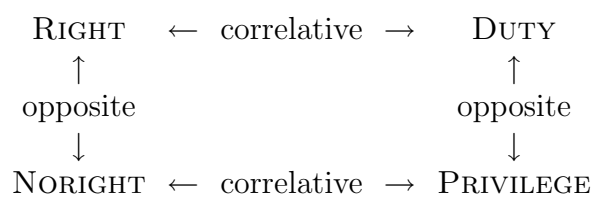
As pointed out in section 4.4 and [244] in general terms, and explained in detail in [243], there is firstly a difference between things one does in one's own interest, and things one does in the public or communal interest, and secondly a difference between powers one can exercise in one's own interest and powers one can only exercise in the public interest.

This discussion of obligation and correlative right extends to Hohfeld's jural relationships, of which duty-right is just one form. Hohfeld's jural relationships can only be understood by appealing to the preferences and intentions that typically come with adopting a certain agent role.

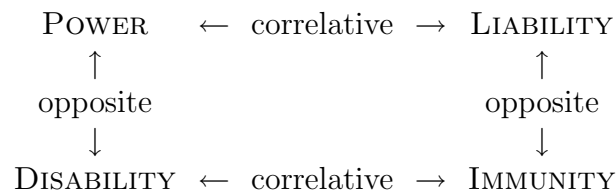
4.8.2. Hohfeld's Conception of Rights and Powers

The principal aim of Hohfeld's work (in [156]) was to clarify *jural relationships* between parties. Hohfeld presents us with an analytical scheme which distinguishes four different categories of jural relationships between parties and makes a number of analytical distinctions between various legal positions.

Hohfeld asserts that there are eight such entities: right, privilege, power, and immunity along with their respective correlates of duty, no-right, liability, and disability. In addition, each fundamental conception is a *jural opposite* to another: privilege, right, power, and immunity are the respective jural opposites of duty, no-right, liability, and disability. These form two squares. The first one is the *deontic square*:



The second one is the *potestative square*:



Hohfeld's system allows for a more fine-grained distinction between patterns we observe in legislation, and generally turns out to be more useful, in particular in private law, for recognizing and classifying patterns in legislative language than the basic deontic categories and constitutive rules in general. It is useful in this context because it covers nearly the whole terrain covered in this chapter, and has been used before in the legal knowledge engineering field (cf. for instance [5]).

Hohfeld conceptualization is not without its critics, however. Hohfeld's names are sometimes slightly confusing. The name privilege for instance suggests more than it actually means. The privilege can also be called plainly permission, right, or liberty. Also right has at times be renamed to claim or claim-right to distinguish it from other uses of the term right. Since section 4.7.3 distinguishes powers or competences from abilities, disability is also an unfortunate name. Ability/disability for "brute" ability, and competence/incompetence (or power/impotence) for institutional ability make more sense in the context of this book.

Also liability is likely to confuse people and has been criticized. It is for instance called subjection in [244]. Hohfeld's liability means that you are exposed to the exercise of a power: you (have to) accept exercise of the power to change legal reality. In private law we encounter a similar concept of the same name, but with a more specific meaning. But rejection of a name like liability for this reason makes it impossible to develop a jurisdiction-independent vocabulary for legal reasoning.

Another problem with Hohfeld's categories is according to some (cf. [135]) that they fail Ockham's test and are therefore not fundamental: there are multiple ways in which his 8 fundamental conceptions can be reduced to each other. This however also applies to the three deontic categories. Knowledge engineers should certainly be willing to take advantage of attempts to interrelate a number of well known concepts from the field of law. There is no need for them to be *fundamental* in Halpin's sense.

Another problem is with the notion of opposite. The opposite pairs seem to have subtly different meanings, which are not quite captured by simply negating the opposite. This is yet another indication that these patterns are hardly *fundamental* legal patterns.

Hohfeld's relationships distinguish between normative positions and other non-normative legal positions, between the competence and incompetence to play a certain agent role, and therefore to cause a certain change of position, and between the obligation to cause a certain change of position or the absence of such an obligation, and most importantly, between the one who acts and the one who predicts the actions of another.

Although we can in principle apply this set of distinctions to any pair of agents, Hohfeld's presumption is that the agent observing and predicting the actions of the

other has an interest in what the other does, and ascribes interests to the other. In essence we are dealing with the ability of one agent to infer 1) that another agent has the ability or inability to change his situation in relevant ways and 2) that the other agent has an interest in changing or not changing it. Hohfeld's concepts of right, duty, privilege, and no-right relate to the interests someone has in changing the situation, presuming they behave normally, and the concepts of power, liability, disability, and immunity relate to the ability to change the situation.

The explanation again has to appeal to scripts (cf. section 4.5): an earlier example given of a "primitive plan" is buying a quart of milk. We realize that the execution of this script depends on someone else's willingness to execute "selling a quart of milk" part, but we still tend to believe that this will be achievable without problems, even for a known reference price. The most obvious reason that we are actively monitoring and predicting someone else's actions that change our situation is because we are executing a script that *involves that other person* as an actor.

Proposition 24. *Jural relationships* only exist between pairs of agents involved in the execution of the same *script*.

Example 11. This has to be made concrete with an example. The initiative to act has to come from both parties, and both parties recognize the same relevant participants of the situation. As an example we will use a sales transaction initiated by an irrevocable offer to sell by y . Person y offers for sale to x a painting for the price of \$500, to be accepted before next Tuesday.

Assume for the purpose of this example, prevailing law is that person y now has a duty to sell, under the stated conditions. Person x therefore has a right to buy under the stated conditions.

We can abstract away the amount (for the price of \$500), the object to be sold, and the timeframe since these do not add to the legal complexity of the case. A not time-limited offer to sell for an unspecified price would however be more difficult to handle. The SALE script involves two actors, the SELLER and the BUYER, who each have the ability to perform two constitutive actions, in no particular order:

$$\text{OFFERFORSALE} \sqsubseteq \exists \text{actor.SELLER} \sqcap \forall \text{actor.SELLER} \quad (4.4)$$

$$\text{SUPPLYSOLDITEM} \sqsubseteq \exists \text{actor.SELLER} \sqcap \forall \text{actor.SELLER} \quad (4.5)$$

$$\text{ACCEPTOFFER} \sqsubseteq \exists \text{actor.BUYER} \sqcap \forall \text{actor.BUYER} \quad (4.6)$$

$$\text{PAY} \sqsubseteq \exists \text{actor.BUYER} \sqcap \forall \text{actor.BUYER} \quad (4.7)$$

From the perspective of the buyer we are dealing with a buy action, and from the perspective of the seller with a sell action.

With this scenario in mind we run through Hohfeld's relations.

Definition 11. A *right* of x towards y wrt. z is equivalent to a *duty* of y towards x wrt. z .

The duty is obviously very similar to the obligation of deontic logic. Party SELLER sets the script into motion by an OFFERFORSALE which is recognized by BUYER, creating an institutional state, or position, OFFEREDFORSALE.

If BUYER proceeds with ACCEPTOFFER, which requires situation OFFEREDFORSALE, he creates for himself, besides the institutional state OFFERACCEPTED, an obligation to pay represented as in section 4.7.3, which is also SELLER's right to receive payment from BUYER. Moreover situation OFFEREDFORSALE and OFFERACCEPTED also creates SELLER's duty to supply the sold item, which is also BUYER's right to the same.

Informally we could also say that the acceptance creates SELLER's duty to "sell the painting", and it is not uncommon to hear it phrased in this way. This obligation is violated if SELLER does not sell the painting. Note however that this duty is also violated if BUYER does not complete his part of the transaction, which is giving the promised amount of money. BUYER may for instance find out that he has less money than he thought and try to back out of the sale. SELLER can however also refuse to take money, and BUYER can refuse to take the painting: the obligations do only apply to the part of the transaction which the party is able to perform.

This kind of statement therefore means in effect that x has a duty to y to perform z with the assistance of y only *if* and as long as y prefers z to happen. In practice this rarely leads to problems if we are dealing with the kind of obligations that only confer a *potestative right* (usually to go to civil court) on the other party. If a third party (e.g. a police officer) would have such a potestative right, the result would be bizarre.

If you have been offered the painting, and the offer is open to Tuesday, you can say now that you will not buy the painting and by saying so relinquish your right to buy it. In general this is however not the case. This *power* cannot be presumed.

Definition 12. A *privilege* of x towards y wrt. z implies *no right* of y towards x wrt. z .

Privilege suggests an exceptional situation. Privilege suggests permission since its jural opposite is the duty. If, for instance, the BUYER in the script is a minor, he has *no* duty to complete the sales transaction. As already noted in section 4.6, the explicit permission functions as an exception to a more general obligation.

Hohfeld considered the no-right as opposite of the right. I, and others before me, see no benefit in introducing a category for the mere purpose of completing the square. A no-right is simply the denial, the negation, of a right. Person y does not have a right towards x with respect to z , *either* because of the presence of a privilege to that effect or simply because of the absence of a duty to that effect. The SELLER for instance has no right to sell the painting before his offer for sale is accepted by anyone.

Readers might object that the sale is a defined legal act, and that it also takes place if BUYER or SELLER fails to complete his part of the bargain. In other words, as soon as BUYER has accepted the sale takes place. There is no way to get out of the sale. Running away with the painting does not change the fact that the sale did take place: it simply becomes disappropriation of the painting.

The rules regulating sales transactions are in actual law *not* normative, and the example is not at all a plausible representation.

The second square deals with this distinction. It is analogous to the first one, but now we distinguish an institutional LAW:SALE that took place:

$$\text{OFFEREDFORSALE} \sqcap \text{OFFERACCEPTED} \sqsubseteq \exists \text{constitutes.LAW:SALE}$$

Definition 13. A *power* of x towards y wrt. z is equivalent to a *liability* of y towards x wrt. z .

In short, a power is one's ability to alter legal relations. BUYER has the power to enter into a LAW:SALE with SELLER of the painting when he has offered it for sale. Thus, BUYER has the power to bind SELLER, *and* himself, to the legal consequences of the LAW:SALE. SELLER, thus, has a liability, which is correlative to power, in that he is liable to having his legal relations altered.

Definition 14. An *immunity* of x towards y wrt. z implies a *disability* of y towards x wrt. z .

If x has an immunity against y with respect to the sale of the painting, it means that y has no power to change x 's legal position with respect to the sale of the painting. Contrary to the permission the immunity is not generally understood as coming in a weak – implicit – and strong – explicit – variety. Immunity is always stated explicitly. The legislator may have stated that minors are immune to sales transactions. If BUYER is a minor, then accepting the offer does not change the legal position to LAW:SALE. A disability to bring about a legal effect may be because the counterparty is immune to it, but more commonly because some other necessary ingredient is missing.

While liability and right are simply other perspectives on power and duty, privilege and immunity are simple *one of* the possible reasons for no right or no ability.

Power and immunity are central to regulating the relations between state and citizen. For instance, if the constitution states that the state has no power to place me under a duty to ask prior permission for expressing my opinions in writing, then I have an immunity and the state a disability. In common parlance this is also a right. So *right* in common parlance may translate, depending on context, to a:

1. simple permission;
2. immunity;
3. potestative right, or;
4. right correlative to a duty.

The scenario discussed in this section, and its relation with scripts, will be continued in section 6.7.1.

4.8.3. Everyone and Delegation

So far we dealt with cases between a clear party x and y . As a general rule, the parties will have a right to go to a civil court in the illegal situations and the court will search

a remedy or reparation to be made by the party responsible for the violation to the party wronged by the violation.

Most norms do not clearly identify both sides of a transaction, however. When x runs a red light while having a duty not to do so, for instance, x wrongs everyone else present in that traffic situation. Other traffic participants have an abstract right towards x with respect to stopping for the red light, but will generally only be in the position to go to court to find a remedy or reparation if there is a concrete wrong. The right correlative to duty does not imply any potestative right. Still x wronged “everyone” in the abstract, and therefore the state exercises the reaction to the violation of a duty on behalf of the community as a whole. This “delegated” power is usually exercised in criminal law. Person x will be fined by the state if caught running a red light.

Hohfeld’s analysis has often been criticized for not dealing with these problems in identifying the involved parties, and sometimes reference is made to “universally quantified” others (cf. [156]). A categorical distinction between natural persons, private and public legal personalities, the abstract state in general, and “everyone” or the community seems a more promising approach. The problem is one of identifying the various implicit delegations of the right to search remedy or reparation from (for instance deceased) natural persons and from “everyone” to the state and its constituent public legal personalities. Since modern democracies are based on a conceptual model of *delegation* of power by ‘everyone’ to parliament, we can safely assume that government has the power to delegate these rights to its own parts.

Proposition 25. There are *collective entities* that can play the role of agent in the delegation of powers, but have no other ability besides that. Entities of this type for instance occur in constitutional law.

Delegation of powers is quite central to how administrative law conceives of the legal position of the state and its parts in many legal systems, and therefore also to the legislative process itself. It is considered a central feature of the conceptual organization of legislation in [51], and is discussed at some points in chapter 5, specifically section 5.2.5.

4.9. Conclusions

Law consists of institutions, and these institutions can be conceptualized as systems with a well-defined interface with an environment. The structures of the institution are defined by the institutional facts that make up the institution, and its mechanisms of change are the constitutive rules that specify what *constitutes*, or counts as, an institutional fact. Conversely, the institutional fact has a constitution base, which consists of the application of a constitutive rule to the *constituting facts*, which are *brute* facts, yielding an institutional fact.

Brute facts are pre-existing and external to the institutional reality constituted by the rules. They form the relevant environment. The institutional facts of one institution can be the brute facts of another one.

This notion of the law as a set of socially recognized systems with an interface opens up the possibility of maintaining a one-to-one correspondence between the knowledge components of chapter 3 and the institution. This has as a consequence that defeasibility can only occur on the interface. As explained in previous chapters, this is justifiable as a form of ontological stratification.

Let C_b be a concept from brute reality, and C_i a concept from institutional reality. *Institutional rules* are terminological axioms, of the form $C_i \sqsubseteq C_i$ or $C_i \equiv C_i$, about the institution and its structure.

Constitutive rules take the form of an indicator or a requirement, and sometimes both forms exist for the same facts. An indicator is a rule of the following form:

$$\begin{aligned} &(\text{DEFAULT } r \text{ (known.} C_b \text{) (free.} \exists \text{law:constitutes.} C_i \text{)} \\ &\quad \text{(assume.} \exists \text{law:constitutes.} C_i \text{)}) \end{aligned}$$

A requirement is a rule of the following form:

$$\exists \text{law:constitutes.} C_i \sqsubseteq C_b$$

The combination of both results in a peculiarly asymmetric kind of equivalence statement.

In addition I have used the integrity constraint to check the integrity of arguments for a certain proposition. The constraint can be used to model the burden of proof of the user in a specific KBS application in a simple way, and has the following form:

$$(\text{CONSTRAINT } r \text{ (known.} \exists \text{law:constitutes.} C_i \text{)})$$

Terminological axioms ($C_b \sqsubseteq C_b$), constraints, or default rules about brute reality do not in principle belong to the legal institution, and do not occur in the source of law: the representation of a rule in the source of law confirms institutional status for at least one concept in the rule.

The identification of institutions with ontological strata of course has consequences for the identification of institutions. Since I use my own criterium for telling apart different institutions (apparent defeasibility is an argument for splitting an institution into multiple institutions), we cannot maintain other such criteria (for instance that each legislator manages its own institutional reality).

Normative rules are also constitutive rules that derive the institutional fact either the *disallows*(n, i) or *allows*(n, i) from constituting facts about i .

The institutional interpretation is a coherent description of the meaning of law that can be used to model knowledge about the sources of law. Note however that institutional and constitutive rules *are not* types of legal rule: the relation is slightly more complicated, as chapter 5 will show.

The institutional interpretation however tells us little about the functions of law for its users. It applies to chess games as well as to law. The purpose of chess games is winning, but the purpose of law is to formalize a normative order. Law makes social interaction predictable by giving people reasons to do certain things and to refrain from doing certain things.

To explain these functions, we have to appeal to planning and plan recognition. These areas do not however form part of the field of law, and neither do the sources of law formulate a particularly coherent ontology of these, in essence, mental and therefore not directly observable activities.

In some cases such an explanation is straightforward. The analysis of normative rules in terms of normative positions and obligation, i.e. deontic logic, is such a straightforward abstract theory of behaviour, based on the expectation that people generally avoid the circumstances in which they are liable to be punished.

To explain the normalizing effect of other rules one must ascribe intentions and preferences to agents: People sometimes intentionally try to bring about or avoid certain legal facts. Since cognitive resources for planning are limited, and people share a lot of their knowledge, intentions are both predictable and can be recognized on the basis of observing behaviour.

The legislator sometimes directly appeals to the idea that behaviour, and the changes brought about by behaviour, are the result of a plan. Some rules for instance only have an institutional effect if intended. Normative rules can address not only the changes brought about by somebody's behaviour, but also the intention behind the behaviour. The subject of a normative rule is therefore naturally the action, which links observable behaviour to a plan. The legislator may also occasionally use concepts in rules – like the right correlative to a duty – that make presumptions about the preferences and expectations of a counterparty (that is not himself *addressed* by the rule but identified by his social or legal agent role) in a transaction.

In this chapter we introduced some minimal vocabulary for action as plan execution, for the changes brought about by plans, for agent roles, and for the relation between situation and action. Of particular importance is the notion of execution: the expression $\exists \text{executes}.P$ means that 1) something is an action, and 2) that the action is an attempt at performing P , which is a task – i.e. a plan being executed. Although tasks and actions are different things, the description of both typically uses the same vocabulary.

Since the subject of normative rules is an action, an analysis in terms of normative positions and obligation based on possible worlds must evidently be based on *alternative actions*: the true object of comparison must then obviously be alternative plans that could be, or could have been, executed.

While this chapter discussed the phenomenon of law in separation from the formal sources of law, chapter 5 will focus on the relation between the sources of law, as the physical result of formal constitutive acts on a legal institution, and specific identifiable *legal rules*, about which we express knowledge in the form of the knowledge representation rules of chapter 3. The function of law, formalizing normative order, will return in chapter 6.

This book is based on the proposition that the law and normative order can be completely separated. Chapter 6 will however discuss some instances where this becomes problematic because the user of legal rules must ascribe an intended normative order to the legislator to apply the rules. It is however equally problematic to want to completely include the normative order in a knowledge representation, and certainly in a knowledge representation of sources of law.

Chapter 5

Representation of Sources of Law

5.1. Introduction

To legislate is to perform formal, legal actions: the legislator, by legislating, represents an institutional fact with the intent of creating that institutional fact. The source of law *represents* the rules the legislator creates, but is separate from them: the source of law will still represent the rule when the intention of the legislator to have the rule exist in institutional reality has already disappeared.

The institutional facts that the legislator creates are institutional rules, constitutive rules, and institutional facts required for the correct functioning of these rules. The ulterior function of these rules is to formalize a normative order intended by the legislator.

The legal rules represented by the source of law appeal to two separate realities – institutional reality and brute reality – and perform a mapping from brute reality – the ontological substratum – into institutional reality – the ontological superstratum. The substratum has an existence independent of the rules, while the superstratum is supervenient on the substratum and exists by virtue of recognition of the rules, because people act as if the postulated effects exist.

Through institutional rules, and indirectly through the things it assumes in the constitutive ones, the source of law represents an institutional ontology that describes an institutional reality: it maps out a logical space of possible models of the institution. These rules can be interpreted as terminological axioms.

The legislator makes assumptions about the structure of brute reality, in particular concerning the behaviour of agents. Plans, intentions, choices, preferences, and abilities play a central role in the way legislators structure reality. The source of law does not, normally, posit terminological axioms about the relevant brute reality, as section 4.4.1 convincingly showed: it merely points to some *legally relevant* terms, and assumes that the users will be able to make sense of them.

Since the quality of the mapping between brute and institutional realities is likely to be imperfect, we should assume that institutional facts only exist as long as it is consistent to believe they exist. This is however based on an assumption, being that our beliefs about brute reality are generally speaking deeper entrenched than our beliefs about institutional reality: if they are in conflict, it is institutional reality that should give way.

This chapter focuses on two issues:

1. the logical representation of information about legal rules; and
2. the problems legislators encounter in managing a large body of legal rules, the solutions they have found for these problems, and the consequences this has for knowledge engineers who try to represent the meaning of these rules in logical form.

One form of logical representation will conspicuously not be addressed: the normative rule. Normativity is one of the subjects of chapter 6.

Section 5.4 functions as a summary of the important representation decisions made in the rest of the chapter, and relates these to the notion of knowledge components representing sources of law.

5.1.1. Auxiliary Rules and Facts

Legislators and users of legislation have developed a number of strategies over the ages to deal with the ever increasing complexity of the mapping into institutional reality. These strategies can be classified as design strategies or interpretation strategies (cf. generally [260]).

Design strategies help resolve potential problems in the use of rules that were already anticipated by the legislator. The *lex superior derogat legi inferiori* (higher law overrides lower law) principle of section 6.6 is an important example of a strategy that must be taken into account during the design of legal institutions. When the user of a source of law is confronted with an apparent non sequitur, he can use solutions designed into the law to resolve it. The *lex superior* principle will only be applicable if a hierarchy is built into the system.

Design strategies typically deal with the existence of multiple legislators, even ones who derive their legislative power from delegation by another legislator, that may contradict each other. In addition they deal with the fact that legislators regularly change the rules. In essence, these strategies help distinguish legal institutions and their respective institutional realities, and prioritize between the actions of different legislators operating within the same institutional reality.

Interpretation strategies on the contrary can also be used to resolve arising ambiguities that were clearly not anticipated by the legislator. The *lex specialis derogat legi generali* (specific law overrides general law) and *lex posterior derogat legi priori* (newer law overrides older law) of section 6.6 are typical examples of strategies that appear to spontaneously arise from use of sources of law.

The legislator can choose to organize his rules in accordance with the intended application of these principles if potential ambiguity is anticipated, but these strategies are applied even if we do not believe the result reflects the intentions of the legislator. As such their status as “legal principles” is questionable: one could take the point of view that these strategies merely reflect deeply entrenched mental habits in understanding messages that are generally respected by the legislator for pragmatic reasons.

Design considerations motivate the legislator to add various *auxiliary propositions* to legislation¹. Some of these are rules, others are simple formal constitutive acts postulating an institutional fact or the occurrence of an institutional event that takes its significance from the rules of the institution, for instance *this statute is repealed on January 1st 2009*.

Auxiliary propositions shape and regulate the domain of legislating itself. In the work on *MetaLex*, discussed in section 5.3 (cf. [53, 41]), they are of central importance: relevant *metadata* (cf. section 3.2) is usually found in auxiliary propositions.

Knowledge representation of the source of law can be considered a simple matter of translation of one language into the other, but the relation between the source of law and institutional reality is more tenuous: the propositions in the source of law may be defeated by others, and the sources of law reflect part of the history of an institution, from which its current state can be reconstructed, rather than an institution as such.

¹Compare Dutch *hulpbepaling*.

The relation between sources of law and their knowledge representations is introduced in section 5.2, and sources of law themselves are discussed in more detail in section 5.3.

5.1.2. MetaLex

Publication of legislation, and the development of tools for working with legislation is at the moment still a jurisdiction-specific enterprise, even if it is standardized at the jurisdiction level. What is required is a jurisdiction-independent XML standard that can be used for interchange, but also - maybe more importantly - as a platform for development of generic legal software.

For vendors of legal software this opens up new markets, and for the institutional consumers of legislation in XML it solves an acute problem: how to handle very different XML formats in the same IT infrastructure. Increasing legal convergence between governments in the European Union, and the growing importance of traffic of people, services, goods, and money over borders of jurisdictions has led to an increased need for managing legislation from different sources, even in public bodies and courts. EU tax administrations for instance need access to all VAT regimes of other member countries to correctly apply EU law, and EU civil courts may nowadays for instance be confronted with the need to understand foreign law on labour contracts to decide on cases involving employees with a foreign labour contract choosing domicile in the country where the court has jurisdiction.

Over the last decade, legislators have begun to adopt XML standards for the formal sources of law they manage, and there is even some activity to standardize on a supranational level. Since these legislator's standards however generally speaking have an institutional status, coordination between countries requires cooperation between governments, and this process moves too slowly from a consumers point of view, and for reasons largely irrelevant to the consumer.

This book is inspired by two XML standard proposals dealing with two complementary aspects of electronic legislation – the documents themselves as a carrier, and an institutional reality they represent: *MetaLex* XML and the Legal Knowledge Interchange format (*LKIF*), which standardized legal knowledge representation. LKIF is yet to be submitted as a proposed standard, and this book deviates in important ways from LKIF.

MetaLex XML is well on its way to becoming formal and de facto standard for legislation in XML. MetaLex XML is also closer to the views of the author of this book: an exposition of what MetaLex is and does in relation to legal knowledge representation therefore suffices for managing sources of law. A description of MetaLex is found in section 5.3.

5.2. Sources of Law and Legal Rules

The legal knowledge source par excellence is the written *source of law*. The source of law is a writing that can be, is, was, or presumably will be used to back an argument concerning the existence of a rule in a certain legal system (see section 3.5.1). It is the

result of a legislative act performed with the intent of creating that rule, and functions as evidence of that legislative act.

Here, again, it is important to stress that the message, the representation of the intent of the legal act, *cannot be* the same as the product of the legal act. The main function of the source of law is to function as evidence that certain institutional events really *happened*. [260] notes the problem:

Note a peculiar feature of any reflexive sunset proposition, for example, proposition C in statute S, saying “statute S will expire at time T”. If one were to ask, some time after T, whether S were valid, the answer would clearly be no. But how do we know that? We know by reference to the wording of C. These words bind us even after they expire. Those who act as if S were still valid may be opposed in court on the ground that present law recognizes that S has expired. The sunset proposition itself may be cited as authoritative on the past effectiveness of the repeal of S. But to do so means that we believe the sunset proposition did not completely swallow itself. Statute S, including sunset proposition C, did expire, but somehow a meta-statement about this expiration did not expire. But there was no such meta-statement in S. If the meta-statement is law, then it appears that it expired; and if did not expire it appears that it is not law. But if we insist after time T that valid legal authority can be cited for the invalidity of S, then we are appealing to such a metaphysical mystery.

The expiration of *S*, as an occurrence, remains part of institutional reality, or more accurately its history, even if *C* is no longer (part of) an active source of law. The legislator directly caused the expiration of *S* in institutional reality: the occurrence is not supervenient on the message.

The written source of law is therefore only the physical *evidence* for the creation of the rule, not the rule itself. As observed in section 4.4, the legislator *represents* an institutional fact with the intent of creating that fact, of gaining social recognition (collective acceptance) for it.

Knowledge representation also attempts to represent the institutional fact; The knowledge representation is therefore not a representation of the written source of law as such: they are rather *both* representations of an institutional fact, in the case of sources of law usually rules (see figure 5.1). OWL axioms and default rules are ideally phrased as statements *about* the rule instead of their reifications being the representation of the rule. By doing this we both avoid the complications involved with reifying RDF statements, and commit to the idea of *metadata* as statements about something.

The document that is considered the source of law, and the legal rules and other legal facts it represents, are subject to different ontological criteria for their existence. In the case of documents a number of different levels of existence is distinguished, as explained in section 5.3. On the *item* level, the physical example of the document, the document is a concrete physical object that is created in some time interval and destroyed in some time interval, and exists in between. On more abstract levels – [245] distinguishes the work, expression, and manifestation level – the document comes into existence but it is not entirely clear whether it ever stops existing *if* it ever does so. As long as one physical example of the document exist it can be brought in as evidence of the past of the legal institution. In rare cases even a document that *claims* a certain

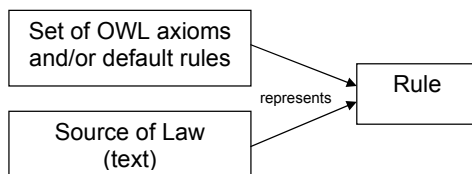


Figure 5.1. Both the original source of law and the knowledge representation are representations of a third object: the rule.

source of law once existed functions as evidence of the existence of a rule².

What is clear is that the existence of the document in no way corresponds with the legal reality it creates. Legislation is usually delivered a considerable time before it becomes applicable, and stops being applicable at some definite point in time.

It is good practice to separate the legislative fact or rule, as an occurrence with a definite starting point and end point in time, from the medium that first – and authoritatively – represented it, without however committing to the equivocation of the institutional rule and its representation as a logical rule: the logical rule that represents a legal rule is also an ontologically distinct entity, created at a different time than the legal rule, and it can be modified during its lifetime even though the legal rule remains the same, for instance because it contains an error or turns out to be a misrepresentation of the legal rule.

Experiences with *MetaLex* (cf. [39], in particular section 6.4.2) have shown that care should be taken to distinguish two senses of applicability constraints on rules, in particular with respect to time and location, as errors are most likely to be made in this respect:

The context of application: Concerns constraints on the time, location, person or agent role, etc., in the action of applying the rule to create an institutional fact.

The matter to which it is applied: Concerns constraints on the time, location, involved person or agent role, etc., in the behaviour to which the rule is applied.

When applying legal rules, one should ask oneself 1) whether the rule can be applicable in the decision one is about to make, and 2) whether it applies to the case, matter, behaviour at hand.

In *MetaLex* the main issue is time and versioning of sources of law. Usually the time interval in which the rule can be applied and the time interval in which the occurrences must have happened to which it is applied are the same, but they may diverge (cf. generally [196, 217, 216, 218]).

Applicability statements, as found for instance in section 6.2, are usually of the second type. The first type has a purely auxiliary function, and is usually of lesser

²Speculation about the laws of ancestors is actually a common theme in medieval sources. The issue became considerably less pressing after book printing became common.

importance in LKBS. Since the context of use of the LKBS is more or less fixed, its representation can often be ignored (for instance the involved agents, for instance a civil administration and its client, remain the same, and the LKBS is deployed only in the time interval in which it can be applied).

The distinction is most obvious in normative rules; If a rule is applicable in a certain time frame, one can apply it in this time frame to determine whether some behaviour you know of is allowed or disallowed. The question whether the behaviour must also have happened during this time frame is however a separate one.

Retroactive applicability for instance normally means that the rule may be applied to occurrences that happened before the rule became applicable. It does not however mean that the rule may be applied before it exists. This means that it is possible that a behaviour is at a later point in time judged disallowed by a court, backed by the existence of a legal rule, even though the legal rule could not possibly have been taken into account when the intention to engage in that behaviour came about.

Ex post facto legislation is regarded as subversive to justice and open to abuse, as remarked by David Hume on the infamous trial of Strafford in 1641:

Better to live under no law at all, and conform ourselves the best we can, to the arbitrary will of a master, than fancy we have a law on which we can rely, and find at last, that this law shall inflict a punishment precedent to the promulgation, and try us by maxims unheard of till the very moment of prosecution.

Prohibition of punishment based on *ex post facto* law is implicit in *nulla poena sine lege* propositions all over the world³. As a general rule, retroactive application of rules that cause harm is considered taboo. On the other hand, one can argue that it is sometimes just to apply a new regulation retroactively because the legislator has become more enlightened in time. It does happen quite regularly in any case.

Similarly, and less contentiously, the law may decide later that events that happened in the past constitute some legal fact now. It is however problematic to ascribe the intent to perform some legal act to behaviour in the past before the existence of the rule that makes the legal act possible, so one does not expect this type of retroaction in relation to legal powers and potestative rights. In the decision to perform a legal act the context of application of the rule and the thing to which it applies are the same.

The notion of *delayed application* is often seen as the mirror image of retroactive application, but is in reality something very different: In this case the maxim is that the rules to be applied are the rules as they were when the occurrence happened. This is a routine technique in law, commonly encountered in routine administrative processes. An operative principle for tax deductibility of premium paid for capital insurance products may for instance be that they are tax deductible if they were tax deductible when the taxpayer entered into the contract.

The result is that there is no single clear-cut time line of valid sets of legal rules and facts that can be used to test cases against: whether a rule is applicable depends both on the events to which it is applied and to the context of application.

³for instance in the Constitution of the United States, art. 1, §9, cl. 3 and §10, cl. 1, or German and Dutch penal codes, both Title 1, art. 1, §1

The world of documents initially seems simpler. Sources of law come in two types: the ones which are not modified during their lifetime (for instance court decisions) and the ones that are (consolidated legislation). The ones which *are* modified by the legislator usually form a neat sequence of versions, only one of which exists at each specific point in time after the initial creation of the work. There are however some exceptions (errata corrige, modification *ex tunc*) which create more than one version of a source of law at the same point in time. These will be discussed in section 5.3.

It is in any case clear that a document management system capable of telling us which version of a document was the one in existence at a specific point in time, does not necessarily help us answer the question which set of rules should be applied to come to a correct decision.

Existent internet portals that make the law available for citizens *are* usually based on the idea that there is a single applicable version of the document for each date; Moreover they may fail to distinguish between the existence of the rule and the existence of the text⁴. It is acknowledged that this timeline may be retroactively changed by the legislator, but the timeline is always clear-cut when one takes a vantage point: If the date of viewing (*sichttag*) is today, then the valid version of a document at some date of interest (the *stichtag*) can be determined.

For the casual user who has a specific context of application in mind this is perhaps the best one can do in a simple user interface. For the knowledge engineer this is not helpful, since the knowledge representation should not reflect a *sichttag*: section 5.3 introduces an event-based method for analyzing the history of a document.

5.2.1. Representation of Legal Rules

Knowledge representation of sources of law consists of formulating logical rules *about* legal rules. The knowledge representation rule is not a direct representation of the legal rule. The central question is whether the special requirements of legislating or legal reasoning tell us something about the relation between the logical form of knowledge representation and the functions of the legal rule the knowledge representation is about.

There are indeed some useful recurring patterns revolving around the **applicable** and **constitutes** properties. The constitutiveness patterns were already addressed in chapter 5 in relation to constitutive rules. Applicability, discussed in the next section, will link the (logical interpretation of the) constitutive rule to the legal rule as an object mentioned in the rule. Constitutive and institutional rules do not need to be strictly distinguished from their logical representation, since we do not need to represent rules about their applicability; Constitutive rules can be applied if it is consistent to do so, and institutional rules must be applied because they are ontological.

A central thesis from chapter 4 is that the distinction of brute reality – the outside not under control of the legislator – and the legal institution – in which the legislator can create his own reality and decide what interfaces for interacting with it exist for

⁴A side note: Dutch Guidelines for Legislative Drafting (*Aanwijzingen voor de Regelgeving*, 168.1) specifically instructs legislative drafters not to realize retroactive applicability with a fictional enactment of the document in the past, to avoid confusion.

others – is paramount. All rules mapping into institutional reality are default rules, because brute reality must take logical priority over institutional reality, and the rest are not, since institutional reality is not intended to be completely arbitrary. Section 4.4.1 showed a worked out example, that however contains some simplifications.

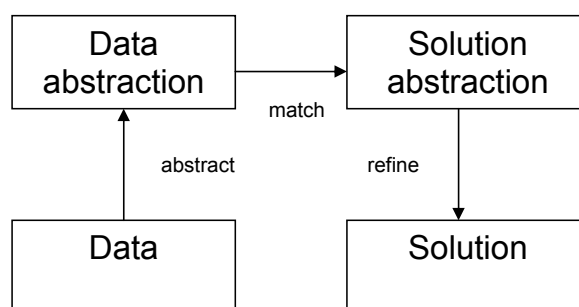


Figure 5.2. Heuristic Classification.

For a functional classification we can find some inspiration in Clancey’s well known horseshoe diagram in figure 5.2 (cf. [81]), which describes a simple problem solving method (cf. section 2.4) originally used for medical diagnosis: We for instance generalize from neutral input like for instance ‘has low white blood count’ to abstractions like ‘immunosuppressed’, and from the abstracted data generate hypotheses like ‘gram-negative infection’, and then flesh this out to a concrete solution by using additional data to distinguish between the subtypes of gram-negative infection arriving eventually at ‘E-coli infection’.

As pointed out in section 2.3.1, abstraction and refinement are not really inverse directions of inference: on the left side of the diagram we abstract *and* aggregate. In law we find analogous classification and reasoning processes. First there is a problem or conflict, which is generalized into legally relevant terms, which *constitute* legal facts, and these help to generate hypotheses about the possible legal instruments available for acting on the problem, and then one checks the details, the necessary conditions: you get hit by a car, which makes you ‘disabled’ and ‘unqualified for work’, which means that you generally speaking are entitled to some social security-related payment, and only then you check out the specific details. Constitutiveness is clearly a form of abstraction, a shift of granularity level of description.

One of the traditional problems with public administration is that it traditionally leaves the responsibility for the first three steps with the citizen, by organizing administration physically in such a way that the citizen has to make the choice to which desk to go. At each desk you can get only one type of ‘solution’, ‘product’, or ‘service’. Without going into the organizational aspects of this phenomenon, this organization does tend to lead to blind spots that affect design, life expectancy, and reusability of knowledge bases and knowledge-based systems.

In [272] we describe a system we developed for the *Juridisch Loket*⁵ in the Nether-

⁵previously *Bureau voor Rechtshulp*

lands, a semi-public organization that gives legal first aid to people with low incomes. The system we developed for this organization handles not only the checklists for determining eligibility for some type of service, but also assists in determining the nature of the legal problem. As we described in [272] one of the more complicated problems is dealing with the fact that many of the clients have multiple problems at the same time (they were unlawfully fired, then divorced, have a dispute concerning the children and the dog, don't have a regular source of income, and are about to be thrown out of their house) and are not capable of separating them in the interview and handling them one at the time.

Besides the special demands this puts on interaction, it also directly suggests a knowledge representation, and a style of legislating. One does not systematically try out all possible checklists. Instead one takes some elements of the story the customers have to tell as *indicators*, and based on the indicators one decides to check which list of *requirements*. The indicator and requirement pattern was also used in section 4.4.1.

The proposition made here is that a lot of legislation is usually organized in this fashion: there are defeasible *indicative rules* that map from a set of indicative conditions to a relevant legal fact. Very often the *intent* of the customer to create this legal fact is assumed by default in LKBS⁶. This is however not always the case: people who have just been fired on the spot are for instance often reluctant to start an appeal to get their job back if the layoff was unlawful. The next step is to systematically check *requirements*, which are necessary conditions. The distinction between indicators and requirements is purely a distinction on the basis of logical form: they are both constitutive rules since they specify a mapping from a brute reality to an institutional one.

Interestingly, the set of requirements applied by the *Juridisch Loket*, which assists people in applying for unemployment benefits, is intentionally less strict than the set of requirements used by the system of the *CWI*, which decides on actual eligibility⁷. The two organizations disagree about the *burden of proof* for certain requirements: the CWI is of the opinion that it is always the applicant's burden, and also has it's own policies for accepting and denying evidence, while the Juridisch Loket takes the point of view that the CWI legally speaking has considerable freedom in these matters, and the applicant in any case has a right to appeal against unfavourable decisions. The LKBS not only implements the *legal rules* the interface shows, it also represents a policy decision of the organization on burden of proof: what information must be supplied before we accept the application? The administrative organization usually even invents its own rules for evidence: the requirement that the applicant must be a subject of the Netherlands is for instance habitually translated into a requirement that the applicant adds a copy of a valid passport by some types of organizations, while others will happily accept verification by oath (i.e. they simply ask).

Compare this to argumentation: indicative rules are used to construct pro arguments, while requirements are used to attack them on the grounds that the conclusion only follows if some necessary condition is met. In this case the burden of proof de-

⁶I.e. "customer wants to create this fact" is not part of the explicit conditions on the adviser's screen.

⁷My own observation.

termines whether it is sufficient to merely question whether the requirement is met, creating a duty for the other side to produce evidence that it is, or whether actual evidence needs to be presented that it is not. In essence the same type of dialog occurs in administrative settings, except that the degrees of freedom are considerably reduced (cf. for instance [53]).

Summarizing, we distinguish between institutional and constitutive rules based on their content, dependant on our view about the boundaries of the institution. Constitutive rules can be further distinguished by the direction of the logical mapping into indicators and requirements. Lastly, we still have the integrity constraints of section 4.4.1 to represent a specific burden of proof assignment to a KBS.

The source of law contains rules of the following logical types (notation following section 3.4):

Institutional rules: Terminological axioms are of the form $r : \alpha \sqsubseteq \beta$ or $r : \alpha \equiv \beta$, where α and β are concepts wholly belonging to the institutional ontology of the institution the legislator intended to create the rule for, and r is the identifier of the rule. Institutional rules are intended to constrain the use of terminology, and are fairly rare in sources of law. Terminological axioms are not intended to be defeasible.

Requirement rules: The purpose of constitutive rules is to specify how institutional reality depends on brute reality. As shown in section 4.4.1, constitutive rules sometimes specify necessary conditions on the mapping between institutional reality and brute reality. Necessary rules are of the form $r : \alpha \sqsubseteq \exists \text{constitutedBy}.\beta$, where α and β are concepts, and α wholly belongs to the institutional ontology of the institution the legislator intended to create the rule for. r is the identifier of the rule. Necessary conditions are not intended to be defeasible.

Indicative rules: Indicative conditions function as the user interface of the institution: they provide the mechanisms by which institutional reality can be changed through action in brute reality. Since brute reality provides the independent variables, indicative rules are treated as defeasible rules, even if they are phrased as definitions. Indicative rules are of the following form, where α and β are concepts, and β wholly belongs to the institutional ontology of the institution the legislator intended to create the rule for:

$$(\text{DEFAULT } r (\text{known } \alpha)(\text{free } \exists \text{constitutes}.\beta)(\text{assume } \exists \text{constitutes}.\beta))$$

Indicative rules are prevalent. Besides their prevalence in sources of law, they are also the category of rules that is easily delegated to lower legislators. The freedom to make contracts also comes down to the freedom to define one's own indicators for certain (already existing) categories of legal facts.

Burden of proof rules: While the previous categories only constrain the valid models of institutional reality, the LKBS is normally based on a specific autoepistemic model (cf. section 3.4.2) that must be settled before a decision follows. In some cases such requirements are explicitly found in legislation. The burden of proof assignment is always a rule in accordance with the possibilities sketched in section 3.4.2, usually a constraint although burden of proof rules may for instance also

“relax” an indicative rule by turning a proposition that must be known into one which may be assumed. Section 4.4.1 shows an example of a constraint.

The `constitutes` (inverse `constitutedBy`) property identifies the boundary of the institution.

Normative rules are distinguished by their (institutional) subject matter – allowed or disallowed – and when represented consist of *both* requirements and indicators. This is the reason not to add them to this list. Section 6.2 presents them in detail. The normative rules are essential for creating normative order: in some fields of law they may be rare, but their presence is essential.

5.2.2. Application of Rules

The autoepistemic burden of proof rule requires judicious use. It by itself does not automatically solve two important *burden of proof* problems for LKBS:

1. As pointed out in section 4.4.1 the burden of proof for necessary conditions usually means that there is an *independent* argument for why the condition is met, that is if the necessary condition is N for C ($C \sqsubseteq N$) there is *another* argument for N besides the trivial and tautological $(\{C, C \sqsubseteq N\}, N)$.
2. If in the current interface of the legal system there are only a few given rules to create legal fact C , the argument for C must be based on one of these rules.

The reconstruction of arguments is not really a direct knowledge representation issue: tracing the reasoning method used in the LKBS will produce the available ingredients for construction arguments for propositions. The second problem is in principle easily solved by explicitly enumerating alternative sets of conditions if you already know all rules that can generate C .

These problems cannot be solved inside a modular knowledge representation of legislation itself: these are exactly the type of assumptions we try to avoid in a context-neutral, reusable, and durable knowledge representation. Consolidated legislation is also usually sketchy on the issue of burden of proof: valid evidence rules are for instance most often found in case law. These issues relate to the epistemic competence of the LKBS (does it know all relevant rules for determining C , or is there reason to believe that some are unknown to the designers of the LKBS?) and policy decisions to be made by the LKBS user (what evidence do we routinely accept for C , and are we open to unforeseen categories of evidence in the main production process?⁸).

LKBS developers will however have a modest requirement for these purposes for a knowledge representation for legislation: it is very useful to know which rules are applied. It is better to state that C must have been the result of the application of r_1 or r_2 , from the point of view of maintenance and compactness, than to repeat the conditions in these rules in a complex constraint rule.

Rules need to be explicitly identified and referred to for several reasons:

⁸The administrative organization may decide, and it often does, that although there may be unforeseen but valid arguments, the appeals process will sort this out.

1. Firstly, if an important function in legal reasoning is to back arguments concerning their existence, they must be explicitly referred to;
2. Secondly, indicative rules (including normative ones) are sometimes subject to choice (cf. section 6.6.3), and must be compared to determine which one to apply. Moreover, if it is not apparent which rule has been chosen we would also not know which one has been applied if no bookkeeping is done; and
3. Lastly, which rules have been applied can be a burden of proof issue: generally speaking institutional facts are only recognized if we know the rule by which they came into existence⁹.

These concerns however apply to the *legal rules* that have been applied and not to the logical rules that represent them. An obvious solution is to do exactly what this explanation suggests: use an explicit representation of arguments based on the format given in section 3.3, and keep track of these during the inference process. If the logical rules are identified by a logical constant r and the represented legal rule by r' , we can state explicitly **represents**(r, r'), and keep track of the fact that ($\{r, \dots\}, C$).

If we however commit to the principle that the logical rules are statements *about* the *legal* rules, we can avoid the reification. Central to this better solution is the **applicable** (inverse **appliesTo**) property, always attached to the object the logical rule is about (i.e. the condition). The following is an example of a representation of a requirement, where C is the institutional qualification and N the necessary condition:

$$C \sqsubseteq \exists \text{constitutedBy}.N \sqcap \exists \text{applicable}.\{r\}$$

The rule states both its effect and the fact that it is applicable. The institutional rule obviously has the same form.

The following does the same with an indicative rule, where I is the indicative condition:

$$\begin{aligned} &(\text{DEFAULT } r \text{ (known}.I)(\text{free}.\exists \text{constitutes}.C \sqcap \exists \text{applicable}.\{r\}) \\ &\quad (\text{assume}.\exists \text{constitutes}.C \sqcap \exists \text{applicable}.\{r\})) \end{aligned}$$

The belief b that institutional fact C only arises through the application of r_1 or r_2 (assume $r_1 \neq r_2$ for any pair of rules) would for instance become, if it were an axiom:

$$C \sqsubseteq \exists \text{constitutedBy}.\exists \text{applicable}.\{r_1, r_2\}$$

Unsurprisingly this rule takes the form of a necessary condition, but it is still not required that it is known whether r_1 or r_2 applies. Only that it is so. This rule format cannot be safely used. A corresponding constraint rule, which does have the required meaning is trivial to reconstruct:

$$\begin{aligned} &(\text{CONSTRAINT } r_2 \text{ (known}.C) \\ &\quad (\text{fails}.(\exists \text{constitutedBy}.\exists \text{applicable}.\{r_1, r_2\}))) \end{aligned}$$

⁹This is however obviously not a terminological truth: we often accept institutional facts simply because we trust the source of information.

This rule says that if C is ever asserted, we must *know* that it was derived from the application of r_1 or r_2 . If the axiom above would have been present we would again get a trivial argument: we know one of the rules was applied because the axiom says so.

5.2.3. Representation of Applicability Rules

The **applicable** (inverse **appliesTo**) property can obviously also be used for applicability restrictions. Applicability statements (cf. [230, 231]), statements of the form *article n is [not] applicable to C* , are similar to the reading of requirements as necessary conditions. The applicability statement is trivially represented (for instance to say that r applies to the acts of civil servants):

$$\{r\} \sqsubseteq \forall \text{appliesTo}.\exists \text{actor}.\text{CIVILSERVANT}$$

Time limiting the application to occurrences is also trivially done:

$$\{r\} \sqsubseteq \forall \text{appliesTo}.\exists \text{causes}\exists \text{extension}.\exists \text{after}.T$$

Similarly, any restrictions on the matter to which a rule or set of rules is applied can be represented as applicability constraints.

A generic format for stating that r_1 states that r_2 (only) applies to C is:

$$\{r_2\} \sqsubseteq \forall \text{appliesTo}.(C \sqcap \exists \text{applicable}.\{r_1\})$$

Slightly more complicated are applicability statements of the form *the rules of act R are [not] applicable to C* , *ministerial directives are only applicable to C* , or *the rules of act R , and rules dependent on it, are [not] applicable to C* .

These involve the grouping of rules into sets of rules. Since the specific applicability constraint on a rule r is based on a corresponding nominal concept $\{r\}$ – the concept that describes the singleton set $\{r\}$ – the logical form of such constraints is also clear:

$$\text{PARTOFR} \sqsubseteq \forall \text{appliesTo}.(C \sqcap \exists \text{applicable}.\text{PARTOFR})$$

PARTOFR is a generalization of a category of rules, all rules part of R :

$$\text{PARTOFR} \equiv \exists \text{partOf}.R$$

When stating applicability restrictions on groups of rules it is however important to keep in mind that the applicability restriction only applies to things of the correct ontological category. If a statute R for instance only “applies to civil servants”, the knowledge engineer has to add himself the restriction that this means “applies to actions of civil servants only where it concerns actions and civil servants only where it concerns actor roles”. A little example for the actions:

$$R \sqsubseteq \forall \text{appliesTo}.(\exists \text{actor}.\text{CIVILSERVANT} \sqcup \neg \text{ACTION}) \sqcap \exists \text{applicable}.R)$$

In reality such broad scoped applicability restrictions often require reading the entire statute, as it is not inconceivable that such a statute that applies to civil servants has some rule that addresses the powers of others towards civil servants.

5.2.4. Also-applicability and Legal Fiction

A problem still exists with a certain formulation of applicability rules: the *also applicable* rule. When there is a rule *n is applicable to C_1* and a rule *n is also applicable to C_2* (which is not a C_1), or alternatively *to be disallowed by n it must be C_1 and C_2 is also disallowed by n*, we have a problem with our reading of the rule as a necessary condition.

In this case we really do need an edit of the first rule to accommodate the second (as suggested in section 4.4.1 as a less preferred solution), or a bizarre *legal fiction*.

Legal fiction is occasionally discussed in legal knowledge engineering literature under the heading of *deeming provisions* (cf. for instance [21, 238]).

Functionally speaking it is generally a rule intended to extend applicability of an existing rule to matters that were clearly not covered by the original rule. If theft is the *taking* of a good, but another rule extends theft to selling something that belongs to someone else, then this rule creates a fiction if according to the original rules theft is necessarily a taking: selling something also becomes a taking, even though this is contrary to common sense and existing requirements.

Since producing evidence to the contrary does not defeat such a rule, the conclusion of a legal fiction can be considered an undefeatable legal presumption. This is contrary to a fundamental principle stated several times: that brute reality takes precedence over institutional reality. As Jeremy Bentham pointed out:

Fictions are to law what fraud is to trade.

To accommodate it within this framework, one needs to clone all rules about THEFT that do apply to the sale as theft, but applied to the new concept SALEASTHEFT, and closely monitor future changes to THEFT to see whether they are also compatible with SALEASTHEFT¹⁰ (cf. for instance [21] for a similar approach, or [238] for another one that does not however address the fundamental problem but only signals its presence).

This representation is no less elegant than the original. *Also-applicability* is a pathological legislator's instrument, particularly when used by a lower legislator to whom the legislative authority to create additional, usually indicative, rules for creating a certain institutional fact has been delegated. The applicability constraint limits the scope of this delegation, and the also-applicable rule bluntly ignores it to expand the scope of applicability. It is a deliberate misclassification. It is occasionally used by courts, and as a legislative technique¹¹.

5.2.5. Purposes of Rules

The identified recurring patterns using the **applicable** and **constitutes** properties can be used with a large variety of different rules. In this section we run by some examples,

¹⁰The legislator may change legislation with the intent of destroying such an established fiction. This may lead to the introduction of new propositions that are idiotic from a common sense point of view, explicitly denying things that make no sense in the first place.

¹¹Legal fiction is prohibited by *Aanwijzingen voor de Regelgeving* 1990, art. 61, in the Netherlands, but this is no guarantee that it isn't done. Strangely, the official explanation of the article then uses also-applicability as an example method of working around the prohibition.

categorized by the purpose the legislator had in mind when creating them.

The legislator believes that by shaping the institution in certain ways certain beneficial effects are brought about in brute reality. The intention of the legislator is generally to improve the normative order: the legislator expresses his own preferences through rules, creating social mechanisms, or at least proposing them, that will have certain effects on behaviour if certain assumptions of the legislator about the preferences and abilities of other agents are met.

The most obvious category of legal rules that has such an effect on normative order is the *normative rule*, which is discussed in section 6.2 in chapter 6, in which the intended normative order will be the topic of discussion. What normative rules do is nothing but attaching the institutional qualifications allowed and disallowed to certain behaviours; The legislator intends us to choose the allowed alternative over the disallowed one. The legislator believes we will do so, because not doing so involves a risk. The allowed and disallowed qualifications perform their function because they are embedded in a larger system of rules that create an incentive for other people to take actions that work out negatively for the agent who violates a norm. This whole system obviously depends on a correct judgment on the preferences of agents addressed by the rules.

The normative rule brings about a legal fact that is evaluated negatively by those who bring it about.

Normative rules are however not the only mechanisms through which the legislator manipulates the normative order in a society. Section 4.4 proposed a classification, based on [244], that distinguishes a number of components that also suggest a function of the rule.

For instance the *intention* to perform a legal act/bring about a legal consequence is a condition for application of the rule:

$$\begin{aligned} \text{RAISEHAND} \sqcap \exists \text{executes.}(\text{RAISEHAND} \sqcap \exists \text{constitutes.BID}) \\ \sqsubseteq \exists \text{constitutes.BID} \end{aligned}$$

This gives people the freedom to choose whether or not they bring about this legal fact.

The legislator may however also restrict the intention with which the legal act may be performed. If T is this intended task, then there is an applicability rule that says that the rule only applies to action performed to execute T :

$$\{r\} \sqsubseteq \forall \text{appliesTo.} \exists \text{executes.} T$$

This means that if one tries to perform the legal act, but without the right intention, then the legal act does not happen. This is not the same as a normative rule which states that such a thing is disallowed. Notions of abuse of power by for instance civil servants depend on normative rules.

Some rules are clearly intended to be used in one's own interest, whenever one likes, but these cannot be distinguished by form. The absence of applicability rules applying to intention is weak evidence for such freedom. One could consider this a power-based

version of the distinction between strong and weak permission, which will be discussed in greater detail in section 6.2: the power is always weak, and its systematic protection can only be organized through non-infringement on it by legislators.

The last recognized category from [244] is the legislator's intent to *enable* people to bring certain things about that could not be brought about in the absence of the rules.

In some cases the legal rule only confers the benefit of legal recognition to actions that would also take place without it (for instance buying and selling), while in other cases the legislator creates a rule solely for the purpose of creating a recognizable way to achieve a certain legal effect (for instance a permit application procedure for gaining permission for constructing a shed in your garden). The legislator sometimes explicitly has to *create* a recognizable way of achieving certain novel legal effects.

In these cases the rules are created in combination with certain other administrative actions: an administrative body or function is set up, a business process is created that implements the intended function, a budget is made available, etc.

The legislator for instance prohibits theft, making theft, which presupposes a victim to whom the good taken belongs, disallowed. The disallowed qualification creates the power, which may be exercised in one's own interest, to report the crime to a body set up for accepting those reports. The body has the power to charge the suspect of the theft with theft before a criminal court, which is to be exercised only in the public interest. The court then has the obligation to determine whether the crime happened, and to penalize the one responsible for the crime, which is supposedly in the interest of the suspect, because it is a foreseeable consequence of reporting the theft. The one who committed the theft generally prefers not to be penalized, and it was foreseeable that committing the theft entailed that risk.

As pointed out by [265], we might think of a creative function of some rules: some rules in a way *create* the courts, public prosecutor, etc, on the institutional level, and there is a quasi-obligation on the legislator that there is also something that *constitutes* it. These are simple propositions that postulate the existence of certain legal facts, with or without a constituting base in brute reality. Postulating the concept, which is by definition non-empty, creates the (possibility of a) court; Declaring instances of it populates the concept:

$$\{c_1, c_2, c_3\} \sqsubseteq \text{COURT}$$

And a judge in court c_1 may for instance be appointed by royal decree, creating an institutional fact, i being the individual:

$$\{i\} \sqsubseteq \exists \text{plays} . \exists \text{memberOf} . \{c_1\}$$

The class of *auxiliary rules* is not identified by its form, but by the regulated domain. Auxiliary rules are identified by the fact that they do not apply to the primary domain regulated by the source of law, but to the context of application of the source of law. To this category legal facts relevant to the legislative domain belong. Auxiliary rules are all those rules that exist because legal institutions have become so

complex that simply publishing a simple list of rules creating the intended normative order no longer works.

Some of these are trivial. In the Netherlands, for instance, formal law declares how it is cited: This act will be cited as the *Act on X* and abbreviated to *X*¹². Other countries simply have a custom that the title is cited, or the date of delivery plus an optional issue number etc.

Choice rules are applicability rules that make the applicability of a legal rule conditional on the applicability of another legal rule. They in other words require that one make a choice between applying two rules, and very often prescribe which one. This makes them normative, but on another level. Representation of choice rules is covered by section 6.6.3. Choice rules are necessary because the legislator cannot guarantee that pairs of rules will never be contradictory.

The source of law may also state *auxiliary facts* that allow one to derive applicability conditions. Instead of stating that *this regulation is applicable from t_1 to t_2 , and may be presumed to apply to occurrences that were initiated between t_1 to t_2* a regulation will typically state *this regulation enters into force on t_1 / is repealed on t_2* , making the meaning of these statements dependent on other law or customs to the effect that a regulation is active and therefore applicable from enactment to repeal. A list of rules would presumably be applicable from assertion to retraction: legislators have however added another layer of complexity by distinguishing the existence of a rule and whether it is active or inactive if it exists. Some legislators allow rules to go into the active state repeatedly.

Explicit delineation of institutional realities may be considered an important auxiliary function, although absence of such rules does not prove that such a delineation does not exist – making all propositions part of the same institutional reality – or alternatively that all sources of law are completely unrelated.

Sources of law may for instance note explicitly that certain concepts are the same concept as the one in other sources of law (e.g. *wages as defined by the act on income taxes and dependent regulations*). Sharing concepts should be distinguished from mere confluence of terminology because of a shared abstraction (every jurisdiction has its own version of *allowed* and *disallowed*, similar in meaning but belonging to different institutional ontologies) or the *deep copying* that we for instance see a lot in insurance policy contracts: every different product has the same definition section copying the same definitions of the same terms verbatim, but they are still different concepts because they can be changed separately if the need arises.

More important for delineation of institutional realities are *mandate*, *submandate*, *delegation*, and *subdelegation* (cf. [38, 289])¹³: this regulates *jurisdiction*, at least in its meaning relating to legislators¹⁴, and which sources of law depend on other sources of law in terms of shared institutional ontology. There is a limited number of legislators whose legislative competence is pre-existent, which includes at least internationally recognized national governments. Their rules de facto apply to a certain groups of

¹²*Aanwijzingen voor de Regelgeving* 1990, art. 184–185

¹³This conceptualization is borrowed from constitutional and administrative doctrine in the Netherlands.

¹⁴I have once proposed to rename this to *imperium*, and to limit *jurisdiction* to the courts.

persons (personam), in a certain territory (locum), and – de facto – in a certain time interval.

Other legislators, subnational and supranational ones, receive their legislative competence from these initial legislators either by mandate or by delegation; The latter means that the original legislator cannot exercise it anymore (i.e. the delegate receives *exclusive jurisdiction* as opposed to *shared jurisdiction*). Furthermore the legislator who receives legislative competence by mandate or delegation may or may not further subdelegate or submandate it.

This structure may be explicit in legislation, which is more likely in civil law jurisdictions, or simply a result of recognition of unwritten customs.¹⁵

Some of the previous patterns for legal acts also apply here: the legislative competence is often constrained by pre-existing applicability rules, and may only be used with a certain intention (for a certain purpose). When the competence is delegated to make rules that create legal facts in the same institutional reality, the rules must obviously also conform to the structure of that institutional reality, as defined by its institutional ontology and other necessary rules: this is why also-applicability (previous section) is a legislative pathology.

For LKBS development auxiliary rules rarely have to be part of the knowledge base explicitly: they are only relevant for determining which sources of law to search for to account for the rules regulating a certain administrative process. When a knowledge representation must however be reusable, jurisdiction, also on matters of substance, must be clearly delineated.

Typically this can be done – without explanation where such constraints came from – with applicability rules:

$$\begin{aligned} \exists \text{initiatedBy}. \exists \text{causedBy}. (\exists \text{actor}. \text{NLGOVERNMENT} \sqcap \forall \text{actor}. \text{NLGOVERNMENT}) \sqsubseteq \\ \forall \text{appliesTo}. \exists \text{actor}. \exists \text{playedBy}. \text{PERSONUNDERNLJURISDICTION} \end{aligned}$$

This example is a gross simplification. Don't take it as a suggestion of how to do this. This type of jurisdiction is firstly determined by the territory (extended to ship and planes choosing that port flag) where an occurrence happened, with some specific extensions to jurisdiction over persons.

The important part is the form of the condition: the legislator simply causes a change that initiates the rule. The properties of the agent that initiated the rule determines the jurisdiction.

The jurisdiction concept, and all complications associated with it, arises because legislators recognize each other's existence. Both the tendency of delegating legislative competence to independent administrative agencies and complex supranational bodies like the European Union tend to complicate the determination of jurisdiction, and lead to an expansion of the role of auxiliary propositions in legislation.

¹⁵A good guide is the constitutional history of the country: bodies that factually existed when the constitutional arrangements of the country took their shape, are typically not created by constitutional law in that country and their legislative competence may be a factual customary one which is recognized.

5.3. Sources of Law and the Semantic Web

MetaLex XML positions itself as an interchange format, a lowest common denominator for other standards, intended not to necessarily replace jurisdiction-specific standards in the publications process but to impose a standardized *view* on documents that function as source of law for the purposes of software development. The MetaLex schema is based on best practices from amongst others the previous versions of the MetaLex schema, the *Akoma Ntoso* schema, and the *Norme in Rete* schema. Other important sources of inspiration are i.a. LexDania, CHLexML, FORMEX, R4eGov, etc. In addition to these government or open standards there are many XML languages for publishing legislation in use by publishers. Standards like PRISM, in which major publishers are involved, are also a source of inspiration.

The MetaLex XML standard recently moved forward significantly, with the adoption of part of it as a CEN¹⁶ prENorm, and its adoption by several industry projects. Many of the participants of the CEN workshop have also been involved in the Legislative XML workshops (see for instance the archive of the front page of the MetaLex web site¹⁷ for previous calls for participation and online proceedings and presentations). In the process of standardization MetaLex changed significantly compared to its previous incarnations (versions up to 1.3.1) for which the author of this book was mainly responsible.

MetaLex is the subject of earlier publications, e.g. [53, 41, 40, 38, 37, 36, 285, 48, 51, 52, 49]. MetaLex is a generic and extensible framework for the XML encoding of the structure of, and metadata about, documents that function as a source of law. It aims to be jurisdiction- and language-neutral, and is based on modern XML publishing concepts like a strict separation between text, markup, and metadata, building on top of structure instead of syntax, accommodation of transformation pipelines and standard APIs, as well as emerging Semantic Web standards like RDF and OWL.

MetaLex, whose first version dates from 2002 (cf. [37]), has been redesigned from scratch in the CEN standardization workshop, taking into account lessons learned from *Norme in Rete*¹⁸ – the Italian standard for legislation – and *Akoma Ntoso*¹⁹ – the Pan-African standard for parliamentary information, and has been submitted as a norm proposal to the CEN.

A partial CEN Workshop Agreement (CWA) now exists. It does not yet constitute a complete, workable XML standard. This partial agreement contains agreements about the abstract content models supported by the standard, the way metadata is added to a document, and a generic model for organizing metadata in RDF. Additional agreements are on the agenda. The description in this section reflects the status of proposals in the summer of 2008, and mainly those features of the standard of interest to and mostly designed by the author of this book.

The *CEN Workshop on an Open XML Interchange Format for Legal and Legislati-*

¹⁶Comité Européenne de Normalisation; European Committee for Standardization; Europäisches Komitee für Normung.

¹⁷<http://www.metalex.eu>

¹⁸<http://www.normeinrete.it/>

¹⁹<http://www.akomantoso.org>

ve Resources (MetaLex), declares, by way of its title, an interest in legal and legislative resources, but the scope statement of the first workshop agreement limits the applicability of the proposed XML standard to sources of law and references to sources of law. As understood by the workshop, the source of law is a writing that can be, is, was, or presumably will be used to back an argument concerning the existence of a constitutive or institutional rule in a certain legal system, or, alternatively, a writing used by a competent legislator to communicate the existence of a constitutive or institutional rule to a certain group of addressees. The scope of the standard covers exactly those things covered by the scope statement in section 3.5.1 of this work.

An important aspect shared by this book and the MetaLex XML standard is the central role of actions and events in interpreting the law. MetaLex however uses another conceptualization of action. Section 5.3.1 explains the relevance of events and actions for a standard for marking up documents.

Identification (section 5.3.3) and reference (section 5.3.4) are also central topics: for the purposes of legal knowledge engineering no functionality is more important than the capacity to make stable and precise links to the things one is representing. The alignment of knowledge bases and sources of law, in particular when it comes to coordinating versioning of both, is the topic of section 5.4.

5.3.1. Metadata and Legislative Occurrences

MetaLex consists of a standard for document structure and identification, and a metadata standard. MetaLex organizes metadata around occurrences – events and actions. For metadata this is uncommon, but it is in effect the same thing as we have done so far with the interpretation of the subject matter of legal rules: these apply to actions, or to the result of actions. There are several good arguments for organizing metadata about legal documents around events and acts, besides a general *descriptive fidelity* argument for law.

Arguments can be characterized as arguments from knowledge engineering strategy, and arguments from legal theory and practice. All are relatively straightforward and unsurprising in this context, but they are nevertheless largely ignored in metadata vocabularies for legislation. A single attribute-value pair, with the document as implicit subject, is often used for such information items as the date of promulgation, instead of reifying the publication/promulgation event and treating the date as an attribute of the event.

A particular metadata description is usually *about* (a snapshot of) some information entity (taken) in a particular *state* – a perceived stability of the entity over a particular time interval that does not take account of changes that are outside the domain of interest. The granularity of that snapshot varies across metadata vocabularies, depending on the targeted community.

This is apparent in the IFLA FRBR conceptualization of bibliographic records (cf. [245]): it groups hierarchically the products of different types of events in the categories work, expression, manifestation, and item. When you make a copy, the item identity changes, but descriptive metadata stays the same. When you add or change metadata statements attached the document, which apply to manifestation, expression, or work,

the manifestation changes, but the expression stays the same, when you edit the text, the expression changes, but the work usually stays the same, etc. When you plagiarize someone else's text, you hardly change the expression, but you do create a whole new work.

To a community that works with certain legislation daily, the insertion of a new provision is for instance an important *event* to be noted, and even to prepare for; For the casual reader it happens to be just one of the many constituting parts of that document's *state* at the moment of consulting.

There are legal theoretic arguments to be made for the importance of event and act descriptions, and the central one is found in the institutional interpretation of the role of legislation (or contracts, or driver's licenses, tax statement forms): One undertakes a legal act on the institutional level by producing a written statement in accordance with a certain procedure. In this reading the document is the mere physical residue of the intentional act that is really important: it functions as physical evidence that a constitutive act that modified institutional reality *happened*, and it declares the *intent* of the act. Evidence is not only found in the central position of legal action and declaration of intent (or will) in legal doctrine, but also in terminology like "Act of Parliament" when one is referring in actuality to the physical *result* of that act of Parliament. It is the act that matters. Of course chapter 4 argued the priority of actions over states in detail.

In the MetaLex CEN workshop (see e.g. [53]) the widely used classification of *event participants* by Judith Dick (cf. [96]) is used: it is interesting to note that Dick developed this vocabulary to describe *legal* text, even though the vocabulary itself is very clearly generic and used in different domains. In this book a more generic format is proposed: Dick's vocabulary has a very functional flavour, and the distinction between types of participants is too arbitrary.

There are several good reasons from the point of view of knowledge engineering, to explicitly reify the events.

One is supplied by Lagoze (see [182]): for establishing semantic interoperability between different metadata vocabularies and for developing mechanisms to translate between them it is only natural to exploit the fact that some types of entities – people, organizations, places, dates, and *events* – are so frequently encountered that they do not fall clearly into the domain of any particular metadata vocabulary but apply across all of them.

It is very clearly the event, or more specifically act, that plays the mediating role between these entities and the resource the metadata description is about. The natural coherence between for instance between *author*, *publication date*, and *publication channel* information (e.g. state gazette bibliographic information) is apparent to all: all are participants in the publication (promulgation) event.

Some other reasons were noted by i.a. the author of this book (cf. generally [51]). Relevant events often transform "input" resources into "output" resources, at the expression or manifestation level, and the respective metadata descriptions for those input and output resources are often the data about the event, i.e. they are shared by the input and output resource: only the perspective is different.

In formal legislation, there is for instance a natural coherence between the old

consolidation, the new consolidation, the modifying legislation, the modifying authority, and the modification date. The modification event, if identified explicitly, links together three different but related resources, and interesting metadata about them.

Different perspectives on this exact same event, because its identity was not made explicit, may yield incompatible metadata descriptions, result in unnecessary duplication of metadata, and several separate occasions in which to make mistakes, therefore unnecessary maintenance, and, lastly, the loss of relevant references between documents. Explicitly identifying events increases the reliability of the metadating process.

As noted in [48], keeping track of changes is especially relevant to law because we have to presume that the law does not become better over time. For most written resources, whether fiction or non-fiction, the last version dominates all others because it is the best: only rarely are we interested in anything other than the current state of the work. We trust that if there is a notable difference at all between today's edition of a book and the first one, today's version will be better. In law we are interested. A tax administration will for instance routinely work at any point in time with at least three different versions: the running tax year, the previous tax year, which is processed now, and the next tax year, which is being prepared. Things like retroactive and delayed applicability, and *ex tunc* and *prospective* versions of legislation, complicate the determination of the current information state about a document; Information on events that happened remains true, and is more easily codified if information is missing.

A version *ex tunc* of a bibliographic work is the product of a fictional legal event that is a temporal fiction of some other event – the constitutive event – that happened *after* the fictional event. The fictional expression creation allows rewriting institutional history, for instance through errata corrige, or at the occasion of an annulment of a modification made in the past by a constitutional court. The fictional event is treated *as if* it were a true event *after* the constitutive event happened. Institutional history therefore appears different depending on which vantage point in time one takes, as already explained (the *sichttag/stichtag* distinction in section 5.2. What one wants to store is however not the snapshot and a description of the vantage point from which one took it, but simply the fact that some brute act caused a legal one *in the past*.

A prospective version also depends on the vantage point, but is conceptually something completely different. It is a manifestation of what is to be expected to be a future version of a work, made when that version doesn't yet exist. Modifying acts are generally published well before the actual modification takes place, in order to give society time to prepare for the changes. It is only natural that early renderings of the future version are made, although there is a risk that the modifying act is retracted or changed in the meantime. Prospective versions may never become law, although they may on occasion still function as evidence that there was at a certain point in time an expectation that it would.

In [51] the point is made that the *expectation* of certain events also functions as a conceptual coat rack for missing information, which was nevertheless essential to the involved organization, a tax administration, in its preparation for future legislation. I refer to that paper for details. We may know that certain events have happened or

expect them to happen, but for instance cannot put a date on it, although we can infer some constraints on it. Essential was in this case that the Uniform Resource Identifier (URI) used in RDF metadata is not a unique bijective identifier: multiple identifiers can refer to the same event (but not vice versa obviously), and what are initially believed to be separate events can – by just stating their equality – be unified without changing the metadata. If we believe that two changes must be made, we can believe in two events, and later merge them into one if it turns out they are made as one action.

In Dick's conceptualization, used in MetaLex, each *occurrent* has one or more participants: Figure 5.3 shows the classification of participants. The *patient* is for instance immanent and product of the action, and undergoes some structural change as a result of the action: at the level of bibliographic entities this applies to the work, while the expression usually takes the role of result or instrument. The instrument is immanent and source of the action, and is not changed during the action: this is for instance the *modifying* expression in a modification of a work, which results in a new consolidation. One of the greater qualities of thematic classification of participants is that it is largely impervious to differences in legal theory.

Consider for instance a Minister of Finance with the competence to index amounts in taxation for the purpose of dealing with inflation. At date t_1 he publishes a directive s_1 to modify income tax law s_2 at date t_2 to compensate for inflation, resulting in s_3 . He uses a specific legislative competence for this purpose delegated to him by s_{comp} . In MetaLex terms we are dealing with the following participants:

The first action:

1. agent: the minister of finance;
2. date: t_1 ;
3. result: s_1 ;
4. instrument: the competence based in s_{comp} .

A background action:

1. agent: some legislator;
2. date: some time before t_1 ;
3. result: s_{comp} ;

The second action:

1. agent: the minister of finance;
2. date: t_2 ;
3. result: s_3 ;
4. instrument: s_2 ;
5. instrument: s_1 ;
6. instrument: the competence based in s_1 .

Figure 5.4 shows the results relating documents s_1 , s_2 , s_3 to events, dates, and persons. It is possible to replace the second action by an event. The difference is that the minister of finance (as an office) no longer has to exist at t_2 , which is in this case immaterial. The directive to act in a certain way at a certain time can be violated,

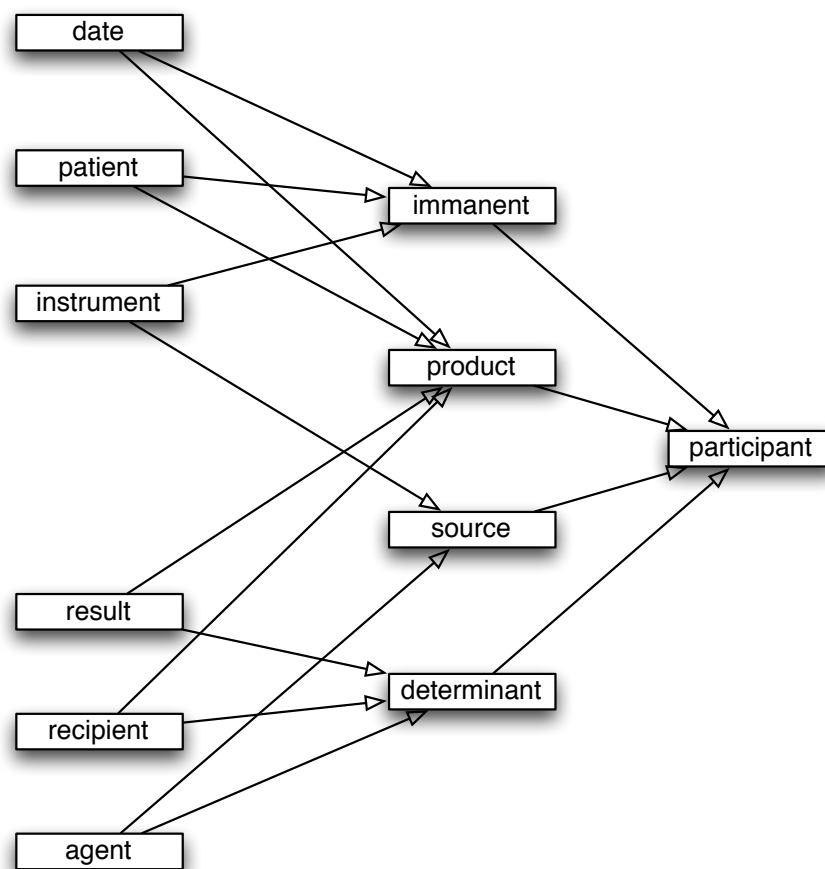


Figure 5.3. Each MetaLex occurrent has one or more participants. The figure shows a taxonomy of participants.

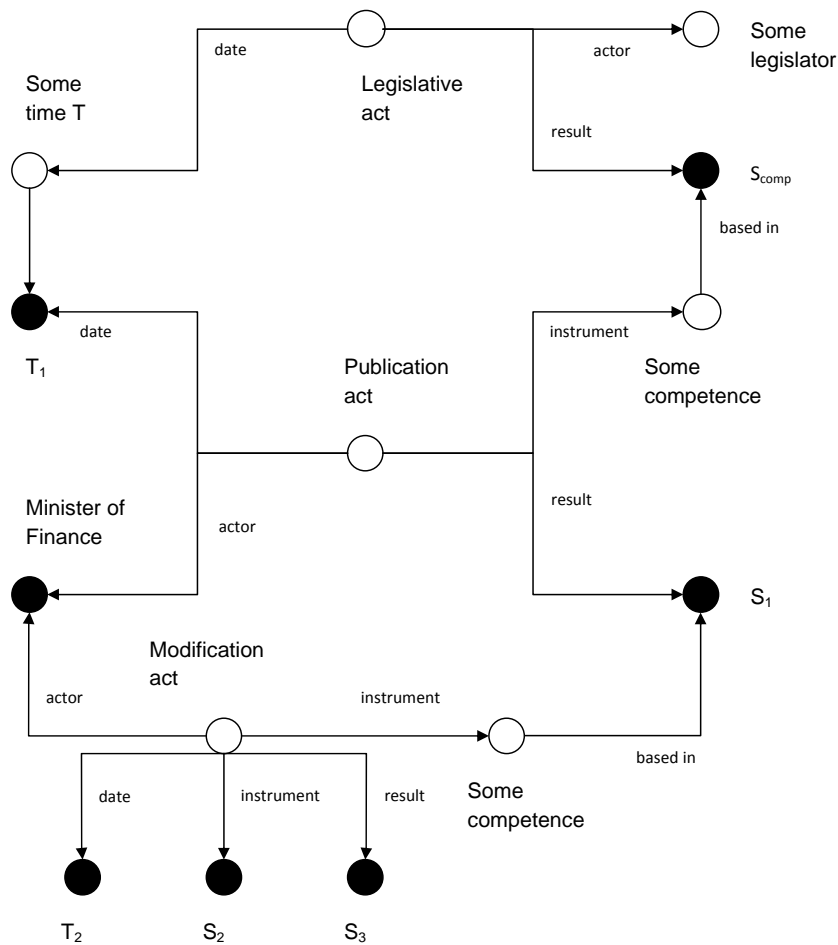


Figure 5.4. Legislative events in MetaLex.

while an event of this type is a purely institutional fact that occurs by definition.

There appears to be a dislike of event descriptions on esthetic grounds in the XML community: they are perceived as in some way less real or objective than dates, persons, and places, and there are obviously no established methods for identifying them, comparable to those for times, places, and persons. This results in reluctance in attaching URIs to these events.

In law, however, it is important and therefore generally clear whether legislative acts (signature, promulgation, enactment, modification, repeal, etc.) happened, and their determination obviously cannot be less objective than for instance the determination of the dates at which they happened.

Organizing information in this way also makes it clear that this legislative domain works in exactly the same fashion as buying and selling, or traffic, or other domains of law. Documents are manipulated in brute reality, while legal rules and facts are manipulated in institutional reality.

5.3.2. MetaLex XML

The use of bibliographic terminology in the MetaLex standard is inspired by the *IFLA Functional Requirements for Bibliographic Records* (cf. [245]):

- A *bibliographic object* is a bounded representation of a body of information, designed with the intent to communicate, preserved in a form independent of a sender or receiver. A bibliographic work, expression, manifestation, and item are bibliographic objects.
- A *bibliographic citation* is a representation of a bibliographic identifier of a bibliographic object, with the intent of referring to that bibliographic object. *Article 1*, *the first article* and *the previous article* are examples of citation, and *the Minister*, *the President of the Republic*, *the accused*, and *We, Beatrix* are examples of references to other, interesting but non-bibliographic, things.
- A *unique bibliographic identifier* identifies a bibliographic object uniquely. The *uniform resource identifier* is used as a unique bibliographic identifier in MetaLex.
- A *bibliographic work* is a bibliographic object, realized by one or more expressions, and created by one or more persons in a single creative process ending in a publication event. A work has an author or authors, and is the result of a publication event. We recognize the work through individual expressions of the work, but the work itself exists only in the commonality of *content* between and among the various expressions of the work: it is an intentional object²⁰.
- An *bibliographic expression* is a realization of one bibliographic work in the form of signs, words, sentences, paragraphs, etc. by the author of that work. Physical form aspects, as typeface or page-layout, are generally speaking excluded from the expression level. Any change in *content* constitutes a gives rise to a new expression. If an expression is revised or modified, the resulting expression is considered to be a new expression, no matter how minor the modification may be. Expression is an intention object.
- A *bibliographic manifestation* embodies one expression of one bibliographic work. The boundaries between one manifestation and another are drawn on the basis of both content and physical form. When the production process involves changes in physical form the resulting product is considered a new manifestation. Thus, a specific XML representation, a PDF file (as generated by printing into PDF a specific Word file with a specific PDF distiller), a printed booklet, all represent different manifestations of the same expression of a work. Manifestation is an intention object. A MetaLex XML element is a bibliographic manifestation.
- A *bibliographic item* exemplifies one manifestation of one expression of one work: a specific copy of a book on a specific shelf in a library, a file stored on a computer in a specific location, etc. Items stored on a computer can be easily copied to another location, resulting in another item, but the same manifestation. This makes adding metadata about the item to the item in principle impossible. On

²⁰I.e. it exists only as the object of one's thoughts and communication acts, and not as a physical object.

the Internet generally speaking only the *uniform resource locator* (URL) is an item-specific datum. An item is a physical object.

Figure 5.5 shows the relationships between the four levels of ontological stratification for bibliographic objects. A MetaLex XML document is a standard manifestation of a bibliographic expression of a source of law. Editing the MetaLex XML markup and metadata of the XML document changes the manifestation of an expression. Changing the marked up text changes the expression embodied by the manifestation. Copying an example of the MetaLex XML document creates a new item.

Work, expression, and manifestation are intentional objects, i.e. they exist only as the object of one's thoughts and communication acts, and not as a physical object. An item is a physical object. Note however that items stored on a computer can be easily copied to another location, resulting in another item, but still an instance of the same manifestation. This makes adding metadata about the item *to* the item in principle impossible. On the Internet generally speaking only the *uniform resource locator* (URL) is an item-specific datum. The item level is therefore not very relevant to XML standards.

The proposed standard is primarily concerned with identification of legal bibliographic entities on the basis of literal content, i.e. on the expression level, and prescribes a single standard manifestation of an expression in XML. Different expressions can be versions or variants of the same work. In addition there is the aspect of role, that relates the bibliographic entity to specific contexts of use: this is consistently treated as metadata.

A MetaLex XML element is characterized by a name, a content model, and zero or more attributes. These are the fundamental content models of MetaLex:

container a container of a sequence of other elements;

hcontainer a hierarchical container of nested elements with titles and numbers;

block the largest structure where text and inline elements mix freely, e.g., paragraphs and other (usually vertically-organized) containers of both text and smaller structures;

inline an inline container of text and other inline elements (e.g., bold); and

milestone an empty element that can be found in the text (as opposed to **meta**).

The philosophy behind content models is explained in [276], and i.a. [53] in the context of MetaLex.

Conformance in the strict sense means 1) validation of XML documents against a schema that includes the MetaLex XML schema, 2) the theoretical possibility of obtaining an XML document that uses solely MetaLex generic elements and validates against the MetaLex XML schema by way of simple substitution, and 3) conformance to the MetaLex CWA written guidelines. Any XML encoding is *transformation conformant* if instances can be transformed automatically into conformant MetaLex XML instances.

The process of declaring a concrete element conforming to the MetaLex norm works as follows:

1. You must use one of the abstract content models for the element;

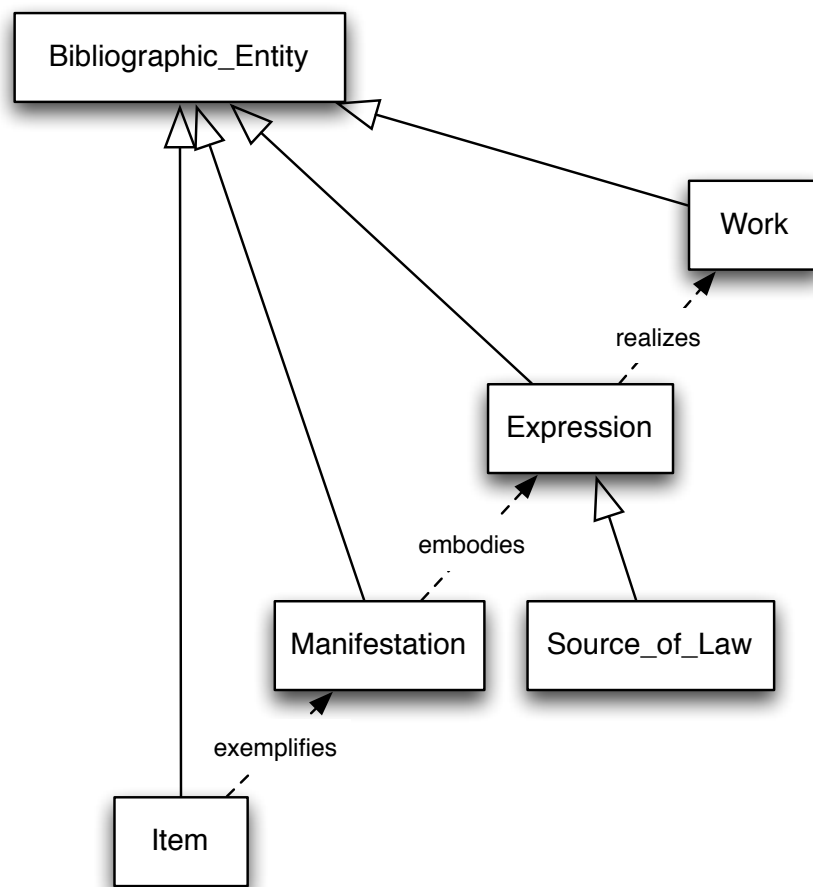


Figure 5.5. A taxonomy of bibliographic entities in MetaLex.

2. You may define a restriction of the corresponding concrete type;
3. You may not define an extension to the content model of a concrete type;
4. You may define an extension of a concrete type for the purpose of adding attributes;
5. You must define the elements as a substitution group of one of the abstract elements and you must identify a type which is either one of the provided concrete types, or the restriction of the content model or extension of attributes of a concrete type that you have defined.

To easily define an element conforming to the standard that can be used in XML manifestations of sources of law, define a non-abstract complex type, for instance a restriction `articleType` of MetaLex type `hcontainerType`, and create an element belonging to the substitution group of one of the abstract elements implementing the MetaLex type specified, for instance:

```
<xsd:element name="article" substitutionGroup="e:abs-hcontainer"
  type="articleType" />
```

Existing vocabularies can usually be redefined in terms of MetaLex content types. It is not sensible to give an example of a MetaLex XML instance here because no such notion exists: MetaLex is intended as a metaschema for other schemas that define concrete XML vocabulary.

MetaLex prescribes what counts as a MetaLex metadata statement, how it is stored inside a MetaLex document, and what classes of entities and which predicates (properties) MetaLex distinguishes: its *ontology*. The RDF ontology is of course extensible. The ontology classifies:

bibliographic entities: the work, expression, manifestation, and item level, and content models;

reference: type of reference between bibliographic entities;

activities: actions and thematic links, and thematic roles of bibliographic entities in at least the actions creation, enactment, repeal;

agent and competence: the agents and institutional instruments (legislative power, etc.) used in legislative activity.

MetaLex uses the conventions of RDF/A processing for embedding RDF metadata statements inside MetaLex XML²¹.

The purpose of MetaLex embedded metadata is nothing more than storage of RDF formatted metadata in MetaLex XML. An RDF description of a resource consists of a set of statements. The MetaLex standard includes an OWL schema that specifies commonly required properties and classes in RDF statements about legal and legislative resources. This schema may be used with RDF stored outside the document in question, and the embedded metadata processing mechanism may be used with other metadata schemas like Dublin Core or PRISM.

The main difference between storage inside and outside the standard XML manifestation is the identification of the metadata author: the metadata inside the document

²¹Consult <http://www.w3.org/TR/xhtml1-rdfa-primer/> for RDF/A.

is associated to the editor of the manifestation, who may be presumed to be the author of the metadata.

A MetaLex document **must** declare what it is a manifestation of, as follows:

```
<meta id="m1" about="" rel="metalex-owl:exemplifies"
      href="/tv/act/2004-02-13/2/tv">
```

Other metadata *may* be embedded.

RDF/A statements *may* be added to any MetaLex element if the content model allows it. Elements derived from the **metalex:urMetaType** type **must** contain RDF/A attributes expressing an RDF statement. Relative URI references in RDF/A attributes are relative to the *xml base* of the containing element.

An RDF/A element is any XML element that contains either the attribute **property**, **rel**, or **rev**. Exactly one RDF statement is generated per **rel** (relation), **property**, or **rev** (reverse) attribute by an RDF/A processor: the attribute indicates a new statement whose *predicate* is the URI value of that attribute. In the case of **rel** and **property**, the subject of the statement is decided by *subject resolution*.

In the case of **rel**, the object is decided by *URI reference object resolution*. In the case of **property**, the object is decided by *literal object resolution*. In the case of **rev**, the subject of the triple is decided by *URI reference object resolution* and the object of the triple is decided by *subject resolution*. If both **rel** and **rev** attributes are used within the same element, two RDF statements are generated.

Literal object resolution yields either the value of the **content** attribute or, if it is absent, the element content. The value of the content attribute is by default interpreted as a plain literal. The element content is by default interpreted as an XML literal.

The **datatype** attribute is used to specify a specific XML Schema datatype (cf. <http://www.w3.org/TR/xmlschema11-2/>). If it is present, the value yielded by literal object resolution is interpreted as an instance of the XML Schema datatype.

URI reference object resolution yields either the URI value of the **resource** attribute or, if absent, the **href** or **src** attribute. The **resource** attribute is only used to specifically communicate that the URI is not intended to be “clickable”, or if a **href** or **src** attribute is already present on the element for other reasons and does not refer to the intended object. It is strongly advised to use the **href** attribute whenever reasonable.

Subject resolution usually yields the URI value of the **about** attribute, or, if the RDFa element that includes the predicate attribute does not have an **about** attribute, the **about** attribute of the first ancestor element that has an **about** attribute. In the absence of an **about** attribute within scope, it yields the *xml base* of the element.

If YYYY-MM-DD is the date of the MetaLex agreement, and one correctly sets the mime type, the associated OWL schema can be found at:

```
GET http://www.metalex.eu/metalex/YYYY-MM-DD
Accept: application/rdf+xml
```

A description logic syntax rendering is found in an appendix of [39] (see also appendix A of this book). The OWL schema specifies the concepts mentioned in section 5.3.3, and the citation metadata specified in section 5.3.4, as well as some related auxiliary concepts.

5.3.3. Individuation and Identification of Sources of Law

In MetaLex, bibliographic entities are identified with URIs or URI references. Identification of bibliographic entities plays a role in:

1. Self-identification of documents;
2. Citation of other documents;
3. Inclusion of document components.

A fourth important purpose for knowledge engineering is that these URIs function as stable target for **represents** links between source and the institutional objects – legal rules, facts – we assume to be represented by them.

Note that URI references that are *relative* can resolve to different URIs dependent on XML base processing context.

Each bibliographic *item* encoded in the MetaLex CEN standard must have at least one URI. Manifestations must refer to the item URI by the URI reference "" (i.e. the empty string URI reference²²). It is in principle not possible to encode item level information in the manifestation. It must be possible to establish the *xml base* of an item, in conformance with the XML Base specification²³ and IETF RFC 3986 (or 2396). The concatenation of the established *xml base* and the **id** attribute of an element **must** result in a valid URI reference for the element, conformant to the addressing recommendations of W3C, which counts as a bibliographic identifier of the element as a bibliographic *item*. There are no further restrictions on item identification.

Note that if one uses an explicit **xml:base** attribute it remains the same after copying the document, which means that it *also* behaves as a manifestation level identifier. There are legitimate use cases of the **xml:base**, where it is inserted as a temporary identifier to an XML subtree in an XML processing pipeline.

The manifestation, expression, and work must also have at least one URI, which counts as their manifestation, expression, work level base, respectively.

Every MetaLex element must have an **id** attribute, not necessarily the **metalex:id** attribute. The **id** value of an element is a manifestation fragment identifier. Concatenated to the manifestation level document identifier it globally identifies the element at the manifestation level, concatenated to the expression level document identifier it globally identifies the expression embodied by the content of the element, concatenated to the work level document identifier it presumably globally identifies a structural element common to various expressions of the work.

Embedded metadata explicitly indicates its subject: if it is the document self-identifier (the empty string) "" or a fragment identifier "#f" composed of the document

²²Note that URI, which is absolute, and URI reference (cf. IETF 3986), which is absolute or relative, and can therefore be empty, are different. URI are globally unique, but URI references are not: only after resolution to a URI they are globally unique.

²³<http://www.w3.org/TR/REC-xml/>

self-identifier and the value *f* of the *id* attribute of the intended target element, then the metadatum refers to the item. The only item-level metadata property is however the *metalex-owl:exemplifies* property, which takes the manifestation-level identifier as value (cf. section 5.3.3). The required type of subject of a metadata property is explicitly constrained in the OWL schema by a domain restriction.

XML attribute values by default pertain to the expression embodied by the manifestation, unless explicitly specified otherwise in this document.

Each bibliographic item exemplifies exactly one manifestation that embodies exactly one expression that realizes exactly one work. Because all these mappings are *functional*, i.e. unambiguously maps to one entity, item identity can be, and often is, used as an indirect identifier of the other objects, similar to how, for instance, email addresses usually have a functional mapping to persons and can be used as an indirect identifier of persons. One can for instance refer to a work by referring to its initial expression in a context where a reference to a work is expected.

The inverse of these relations is however often not a function. One can think of the work as an abstraction of 1+ expressions, the expression as an abstraction of 1+ manifestations, the manifestation as an abstraction of 1+ items. The manifestation, expression, and work are intentional objects whose existence is conditioned to the existence of at least one item, manifestation, expression, respectively. There is normally speaking no such thing as an expression that is not embodied, a work that is not realized, etc.

Besides the hierarchical constitutive relationships between the four levels, there are also horizontal relations between the objects within a level. The expressions of a work in the legal field are usually either the initially published expression, or expressions derived by content modification activity or translation activity. Manifestations of an expression are either the initially created one(s), or manifestations derived by editing activity. Items of a manifestation are either the initially created one(s), or copies of them.

The MetaLex OWL schema includes a number of event type definitions to make these horizontal relations explicit for the expression level. At the item level they cannot be embedded in a MetaLex item for obvious reasons.

Appropriate manifestation level version management methods and tools already exist (CVS, SVN, etc.). These however, from the point of view of the source of law, merely do manifestation-level version management, while most legislative events happen at the work and expression levels. Content-related events like markup, metadating, and digital signature happen at the manifestation level.

The MetaLex standard aims to provide metadata for describing both the hierarchical and relational way of positioning and identifying bibliographic objects, at least at the work and expression level, at the relevant levels of granularity.

Some additional terminology captures some bibliographic phenomena relevant to law, as well as operational criteria for recognizing these in a MetaLex XML document. Their specifications are found in the MetaLex OWL schema:

- A **bibliographic source of law** is a bibliographic object that can be, is, was, or presumably will be referred to, by way of bibliographic citation, to back an argument claiming the existence of a legal rule in a certain legal system, or, al-

ternatively, a bibliographic object published or realized by a competent legislator to communicate a legal rule to a certain group of addressees. Both the legislator and the user of the bibliographic source of law understand it as a medium used for communicating the existence of legal rules, including auxiliary declarations required for the proper understanding of legal rules, between legislator and user.

- An **initial version** of a bibliographic work is the expression that realizes the work at the time of its official release in the public domain as a bibliographic object. It is the `metalex-owl:result` of a `METALEX-OWL:INITIALBIBLIOGRAPHIC-CREATION`.
- A **version** of a bibliographic work is either the initial version of the work, or an expression realized by modification of a version. It is the `metalex-owl:result` of a `METALEX-OWL:BIBLIOGRAPHICMODIFICATION` of the (`metalex-owl:matter`) previous version.
- A **version in force** of a bibliographic work is a version that is, was, or will be *in force* during a specific time interval. The in force time intervals of versions in force of the same work do not appear to overlap in time *when viewed from any specific vantage point in time*. Note however the possibility of *virtual* and *ex tunc* expressions, which may create a substantial difference between vantage points in time.
- A **version ex tunc** of a bibliographic work is the `metalex-owl:result` of a `METALEX-OWL:FICTIONALEXPRESSIONCREATION`, which is a temporal fiction of (`metalex-owl:temporalFictionOf`) some other event – the constitutive event – that happened *after* the fictional event. The fictional expression creation allows rewriting institutional history, for instance through errata corrige, or at the occasion of an annulment of a modification made in the past by a constitutional court. The fictional event is treated *as if* it were a true event *after* the constitutive event happened. Institutional history therefore appears different depending on which vantage point in time one takes.
- A **consolidation** of a bibliographic source of law is a version realized by the execution of legal rules found in another bibliographic source of law to the previous version. It is the `metalex-owl:result` of a `METALEX-OWL:LEGISLATIVE-MODIFICATION` of the (`metalex-owl:matter`), which is the previous version, by the (`metalex-owl:instrument`), which is the other bibliographic source of law stipulating the modification. This is often when it enters into force, although the modification may of course be conditional upon some other event.
- A **variant** of a version in force of a bibliographic work is an expression that shares its *in force* time interval. Although the concept variant is often nominalized, it is a symmetric relationship (`metalex-owl:variant`) between two expressions.
- A **language variant** of an expression of a bibliographic work is an expression that shares its in force time interval, and differs in nothing but language. For instance, the English, Dutch, Italian, and German versions of a European directive are different language variants. It is also a symmetric relationship (derived on `metalex-owl:variant`). In a MetaLex manifestation of the expression this is expressed in different values of the `xml:lang` attribute.
- A **translation** of an expression of a bibliographic work is an expression that

shares its in force time interval, differs in nothing but language, and has been realized by way of translation of one expression into another expression. Translation is an asymmetric relation between bibliographic expressions, expressed by a METALEX-OWL:TRANSLATION event, which has a `metalex-owl:translator`, taking the initial expression as a `metalex-owl:matter`, and the translated document as `metalex-owl:result`. The translation should not be confused with language variant: while language variants can be realized concurrently by the legislator, and are equally authoritative if they are, the translation of an expression is generally speaking less authoritative than the expression it is a translation of, even if *officially* translated.

The list above provides a number of interesting relationships between sources of law mediated by events. For the purposes of explanation of both the use of metadata in MetaLex, and the minimal identifying set of information, we focus on the identifying set as determined by the naming convention. This set is obviously normative, and not descriptive, and aims at URI references that:

1. can uniquely identify sources of law, regardless of jurisdiction and legislative technique;
2. can be reconstructed with a high degree of intercoder reliability; and
3. do not depend on a vantage point in time.

The expression and the work must be declared in case of non-conformance to the naming convention which is not discussed here. Noting that the URI reference `about=""` refers to the document itself, the following declares a standard manifestation, expression, and work base (using the naming convention):

```
<meta id="m1" about="" rel="metalex-owl:exemplifies"
  href="/tv/act/2004-02-13/2/tv">
<meta id="m2" about="/tv/act/2004-02-13/2/tv"
  rel="metalex-owl:embodies" href="/tv/act/2004-02-13/2">
<meta id="m3" about="/tv/act/2004-02-13/2"
  rel="metalex-owl:realizes" href="/tv/act/2004-02-13">
```

The RDF reading of **m1** is as follows: **m1** is a statement that states that the (referent of) `metalex:exemplifies` of (the referent of) (`empty string`) is (the referent of) `/tv/act/2004-02-13/2/tv`.

Read for `meta` in the examples above any appropriate element that permits metadata attributes. The URIs are relative, in this case conforming to the naming convention: the base is set by the processing environment. This means that an identifier conformant to the naming convention describes a URI *reference* that potentially resolves to large set of URIs for each bibliographic object: one for each processing environment that sets its own base.

Note that if one wants to conform to the naming convention but also want to embed another identifier, this is trivially achieved with a metadata statement of that impact, for instance:

```
<meta id="m" about="/tv/act/2004-02-13/2/tv" rel="owl:sameAs"
      href="http://foo.tv/123456" />
```

This simply states in OWL that $c_1 = c_2$, where c_1 is `/tv/act/2004-02-13/2/tv` and c_2 is `http://foo.tv/123456`.

If the naming convention is not used, a set of metadata must be available, either in the form of RDF/A statements, or in the form of RDF. . Let `<<ManifURI>>` be the identifier of the manifestation, `<<ExpURI>>` be the identifier of the Expression, `<<WorkURI>>` be the identifier of the work, all declared as in the previous section. Datatyped values `<xsd:date>`, etc. are values conforming to the XML Schema datatype `xsd:date`, etc. The relevant set of metadata for the work is the following:

```
<meta id="w1" about="<<WorkURI>>"
      rel="metalex-owl:resultOf" href="<<WorkCreationEventURI>>"/>
<meta id="w2" about="<<WorkURI>>"
      property="metalex-owl:workClassID" content="<xsd:string>"/>
<meta id="w3" about="<<WorkCreationEventURI>>"
      rel="metalex-owl:date" href="<<CreationEventDateURI>>"/>
<meta id="w4" about="<<WorkCreationEventDateURI>>"
      property="metalex-owl:xsdDate" content="<xsd:date>"/>
<meta id="w5" about="<<WorkCreationEventURI>>"
      rel="metalex-owl:country" href="<<CountryURI>>"/>
<meta id="w6" about="<<CountryURI>>"
      property="metalex-owl:countryCode" content="<xsd:string>"/>
<meta id="w7" about="<<WorkCreationEventURI>>"
      rel="metalex-owl:issue" href="<<IssueURI>>"/>
<meta id="w8" about="IssueURI"
      property="metalex-owl:issueID" content="<xsd:string>"/>
```

The main purpose of the identifying set is to create a non-ambiguous sequence of work creation events for each work class. Work class is identified by a descriptive string (e.g. "act" for act of parliament, or "municipal;foobartown;decree"). It is recommended to identify work classes with specific publication channels that can be monitored. Work classes are specific to countries, identified by the two or three characters from the ISO 3166-1 standard. The combination of the date of the event and the issue number (some string) uniquely positions the work in the temporal sequence of works belonging to that work class.

The relevant set of metadata for the expression is the following:

```
<meta id="e1" about="<<ExpURI>>"
      rel="metalex-owl:resultOf" href="<<ExpCreationEventURI>>"/>
<meta id="e2" about="<<ExpURI>>"
      property="metalex-owl:lang" content="<xsd:language>"/>
<meta id="e3" about="<<ExpCreationEventURI>>"
      rel="metalex-owl:date" href="<<CreationEventDateURI>>"/>
<meta id="e4" about="<<CreationEventDateURI>>"
```

```

    property="metalex-owl:xsdDate" content="<xsd:date>"/>
<meta id="e5" about="<<ExpCreationEventURI>>"
    rel="metalex-owl:issue" href="<<IssueURI>>"/>
<meta id="e6" about="<<IssueURI>>"
    property="metalex-owl:issueID" content="<xsd:string>"/>

```

The <<CreationEventDateURI>> may and will often coincide with the creation of the work in unversioned documents. In case of an *ex tunc* modification the creation event is a *FictionalExpressionCreation*, and it is a temporal fiction of some other event, to be filled in, which happened at another relevant date:

```

<meta id="e7" about="<<ExpCreationEventURI>>"
    rel="metalex-owl:temporalFictionOf" about="<<RealEventURI>>"/>
<meta id="e7" about="<<RealEventURI>>" rel="metalex-owl:date"
    about="<<RealDateURI>>"/>
<meta id="e8" about="<<RealDateURI>>"
    property="metalex-owl:xsdDate" content="<xsd:date>"/>

```

The *temporalFictionOf* property is for us simply a kind of *constitutedBy*. The relevant metadatum for the manifestation is the following:

```

<meta id="m1" about="<<ManifURI>>" property="metalex-owl:fileType"
    content="<xsd:string>"/>

```

This metadatum may be encoded in the XML manifestation as metadata, but it is assumed that extraction of RDF metadata from a standard MetaLex manifestation already provides the information that the file type is *xml*. The OWL schema provides more identifying metadata that **may** be added.

5.3.4. Citation, Inclusion, and Reference

References in MetaLex documents are usually made with the **href** or **src** attribute. Because the href and src attributes are also RDF/A attributes the referring element also encodes a metadatum if an **about** and **rel** attribute are present. Using this form – the external reference as a metadatum – reflects the point of view that resolving citations etc are acts of interpretation of the source.

The MetaLex standard distinguishes three forms of addressing:

1. Reference to external objects (agents, events, etc): these are always encoded as metadata, and if embedded, always using the **href** or **resource** attribute.
2. Citation of other bibliographic objects: this can be achieved through the citations attribute group, or through metadata, always using the **href** attribute. The target should either be identified by use of the naming convention, or by identifying metadata. In some cases the citation can be embedded in both attribute form and embedded metadata, because of the double interpretation of **href**.

3. Component inclusion: this can be achieved through the **src** attribute and either conformance to the naming convention, or additional metadata. The inclusion can also be made into a metadatum.

The distinguishing property of references is that the (**rel**) value is a subproperty of **metalex-owl:refersTo**. Citation uses a property value that is a subproperty of **metalex-owl:cites**, which is itself a subproperty of **metalex-owl:refersTo**.

This means that the fact that some MetaLex XML element refers to another entity can also be stored in RDF, external to the MetaLex XML document.

A *reference* is something that refers to or designates something else, or acts as a stand-in for a relation between two things: the *referrer* and the *referent*. Since a relation can also be identified, the generic form of a reference is (*referrer*, *predicate*, *referent*), where predicate is the name of the relation, and are represented as RDF or RDF/A.

In the sense intended here a reference is an XML element (directly or indirectly) containing text, and the text is deemed to refer to something else. The XML element will typically be of the **inline** content model type. A *citation* is an expression that refers to something intralinguistic, i.e. to another XML element (directly or indirectly) containing text, or to the bibliographic (work or expression) objects directly or indirectly embodied by it. Other references refer to something extralinguistic, i.e. something other than text, recoverable from the context in which the document was produced.

Article 1, *the first article* and *the previous article* are examples of citation, and *the Minister*, *the President of the Republic*, *the accused*, and *We*, *Beatrix*, *etc.* are examples of relevant references to other things. In a more general sense, any term refers to a corresponding concept. It is good practice to let references to individuals refer to a nominal concept (i.e. the singleton set).

A similar distinction, that should be distinguished from the previous one, is the distinction between exophora and endophora in linguistics. Take the following sentences:

1. *Theft* is the unlawful taking of a good wholly or partially belonging to another.
2. *It* (*The theft*) must have been done with the intent to appropriate.

It (or *The theft*) obviously refers back to Theft in the previous sentence. It is an *endophoric proform expression*, as opposed to exophoric (*That* must have been done intentionally, pointing to an act of taking in progress). Exophoric proform expressions are obviously rare in legal resources.

There is however a difference between the reference *It* referring to theft as defined in the previous sentence and *the previous sentence* as referring to a sentence. Theft and it are coreferents of something other than text: it is not an intralinguistic reference. Expressions like it do stand in for another expression (Theft), but only to indirectly reference meaning recoverable from context.

They are used to avoid repetitive expressions and in quantification (i.e. carrying a variable from one sentence into the next one). It is positive evidence that:

- that the sentences refer to the same thing, and therefore share a concept in the

- same institutional reality; and
- maybe even that they may be taken together to represent a single rule.

This only applies to sentences close together. One can maintain the principle that a rule or fact is represented by a single fragment of text, i.e. structure *isomorphism* from text to representation, but the intended fragment would be the immediate container of both sentences. The implied thesis is that the immediate container is in this case the smallest fragment of text that could be retrieved as a justification of a rule: if text contains endophoric proform expressions crossing the border of sentences, the sentences obviously cannot be read separately.

The following is an example of a reference metadatum in RDF/A:

```
<meta about="#x" rel="metalex-owl:refersTo"
href="http://gov.tv/act/2004-02-13/concepts#theft">
```

The use of references is optional. The **src** attribute may not be used.

A *citation* in legislation normally refers to a work, unless explicitly noted otherwise. A reference to legislation in a court verdict necessarily refers to an expression. Citations in other documents can refer to any level.

The reference to a text fragment is taken as evidence that the represented rule is about the rule represented by the other text fragment. Citations in legislation typically do one of the following things:

- restrict (or extend) applicability of another rule;
- when referring to something that is taken as a definition of a term; indicating that a term used here refers to the term as restricted in the other rule;
- justify the present rule; either by indicating where legislative competence comes from, or an obligation to implement something, or the permission to do so, etc.

A reference with a single target (e.g. *article 1 of the Income Tax Law* or *article 15.3*) is easy to mark up. The entire referring text can be marked, and it should be linked to the one concept that is being referred to. This can be either a work or an expression, depending on whether or not version information is specified or hinted at.

References with multiple targets are more complex to mark, as we now have several concepts to link to (i.e. *article 5 and 6 of the Income Tax Law*). An important problem is where to anchor these links. In addition there are references to ranges, such as *articles 5-10 of the Income Tax Law*. This reference will have several targets, and without studying the target document we cannot determine the exact target locations. It might be *article 5*, *article 6*, *article 7*, *article 8*, *article 9*, *article 10* or *article 5*, *article 6*, *article 7*, *article 7a*, *article 7b*, *article 8*, *article 10* or even *article 5*, *article 10*. The exact targets cannot be included in the document, as they cannot be derived from the information present. In addition, the targets may change depending on the version being referred to in case of references to works.

Ranges get even more complex when exceptions are involved, such as *articles 5-10 with the exception of article 9*. An interesting variation occurs when an exception is made to something other than a range: *article 5 with the exception of the first member*. This implicitly turns article 5 into a range of *member 1*, [...] *member n*.

A final format for references is the “each time”: *articles 10, 12, 15 and 16, each time the first member*. First member by itself cannot be a reference to something here. The links may not be to the articles but instead must be to the first member of each article.

The advantage of special markup for exception and each time constructs is that it is a hint to (less refined) parsers that discover and markup references that this one has already been discovered, interpreted, and discarded.

Each of these complications show the same pattern: there is a difference between the entities explicitly cited in the text, and the ones it refers to. In some cases an element must be explicitly cited to convey the information that it is *not* being referred to.

Complex references create two problems for the MetaLex standard:

1. There are multiple reasonable ways to mark up the same statement. Some XML standards for instance use start and end range attributes on an element enclosing both start and end of the range. Other standards have two separate elements marking the start and end, each using the same href (or similar) attribute.
2. The encoding of a metadata about a complex reference will necessarily require more than one metadata statement, and therefore need more than one carrier element if RDF/A is used. This is a problem if the aim is to declare existing XML documents as MetaLex compliant documents without adding new XML elements.

There are two methods for describing references. The following is an example of a citation *metadatum* in RDF/A:

```
<meta about="#x" rel="metalex-owl:cites"
href="http://gov.tv/tv/act/2004-02-13/2/tv#x">
```

Other available properties (`metalex-owl:citeFrom`, `metalex-owl:citeUpto`, `metalex-owl:excluding`, `metalex-owl:excludeFrom`, `metalex-owl:excludeUpto`) can be used for complex references if the information is stored in RDF or if RDF/A is strongly preferred and modifying the XML document is no problem. The **href** attribute **must** be used.

Alternatively the **metalex:citations attribute group** can be used, if adding elements to the original XML file is a problem but embedding the references inside the document is nevertheless strongly preferred, with the attributes `metalex:href`, `metalex:upto`, `metalex:exclude`, `metalex:excludeTo` to specify two types of ranges from `metalex:href` to `metalex:upTo`, and from `metalex:exclude` to `metalex:excludeTo`. These can be automatically translated into RDF statements.

If the naming convention is not used, the target of the citation must be identified in accordance with the principles set down in section 5.3.3. The use of citations is optional.

Some XML manifestations may *include* components by *reference* instead of by physical inclusion as a fragment, i.e. the inclusion reference replaces a part of the MetaLex manifestation stored in an external object identified by a URI.

Any part, except the top level container, of a standard metalex XML manifestation can be implemented as a inclusion reference to an external object. On the manifestation level you make choices about, for instance, object names and media formats (tiff, jpeg, pdf, etc.). In some cases a text that is (or could be) embodied by a metalex manifestation (for instance a Chinese appendix of a treaty) is embodied alternatively by a media object.

For this the **metalex:srcatt attribute group** is provided, which defines the **metalex:src** attribute. Presence of the src attribute includes a component of the MetaLex manifestation stored in an external object identified by the URI that is the value of the src attribute.

While not technically necessary, the existence of components can also be easily describes in the form of RDF/A metadata simply by adding the property **metalex-owl:component** in the **rel** attribute. The following is an example of a component inclusion metadatum in RDF/A, superimposed on the **metalex:src** attribute:

```
<meta id="x" about="#x" rel="metalex-owl:component"
metalex:src="http://gov.tv/tv/act/2004-02-13/2/tv@/table1">
```

If the naming convention is not used, the target of the inclusion must be identified in accordance with the principles set down in section 5.3.3. The **metalex:src** attribute **must** be used if the **metalex-owl:component** property is used.

While reference pertains to non-bibliographic objects, and citation to works or expressions, components are typically manifestation-level objects. None of these mechanisms is however ever used to *locate* items, even though the used URI and URI references have the form of URLs and URL references. The reason for URLs is a pragmatic one: they are easy to obtain from registrars, cheap to lease, and many people have access to a domain they can consider their own.

5.4. Summary and Conclusions

At this point we have assembled all ingredients for a representation of the rules and structures of legal institutions in OWL DL knowledge components, and we have established a detailed mapping between the sources of law as documents and the rules and structures of the legal institution. At this point it is therefore possible to give a presentation of a methodology for representing sources of law in OWL, and maintaining isomorphism between these resources over time.

Note however that the normative order is in no way taken into account.

5.4.1. Knowledge Components and Sources of Law

Central to reusability in legal knowledge engineering is the notion of a knowledge component, in particular the knowledge component representing a source of law (refer to chapter 3). The knowledge component is stored in the form of a set of *logical* sentences, in no particular order. An OWL DL knowledge component contains only OWL axioms, while an autoepistemic knowledge component consists completely of

autoepistemic sentences. Ideally, for the purposes of LKBS maintenance, one maintains a one-to-one mapping between knowledge components and the sources of law it represents, in order to:

1. facilitate maintaining *isomorphism* between source of law and knowledge representation over time, and to
2. be able to make a single change to *all* deployed LKBS that use the same source of law at a single point.

Single point maintenance is based on the assumption that the LKBS that use the component are able to import the new knowledge component whenever it changed, that one has reasonable assurances that the change, if minor, does not introduce logical inconsistency, and that the epistemic competence of the involved LKBS remains intact, or, in the case of major changes, that the effort involved in impact analysis is minimized.

This expectation of localized impact is however based on the assumption that the institution's design rationale for sources of law and the knowledge engineer's design rationale for knowledge components are aligned. One notable difference is that the legislator does not distinguish between an axiomatic part and an autoepistemic part of the source: most sources of law correspond with two knowledge components, and the autoepistemic one always uses the OWL DL one.

The source of law at the expression level is identified by a URI, *refers* to a set of URI-identified *terms* (**metalex:refersTo**), and it *represents* a set of URI-identified *legal rules* and *legal facts* (**metalex:represents**). This applies to the source of law as a whole, and to any URI-identifiable fragment of the expression. Since the expression does not change, one would expect that the set of legal rules and facts it represents and the set of terms it refers to also normally do not.

The referred terms are usually explicitly present, but in text fragments it is often necessary to first recover the targets of endophoric proform expressions from the context of discourse; This is clearly a representation step, which involves a choice between viable candidates from a linguistic point of view. In addition, a source of law may use two terms, that are the same from a linguistic point of view, but turn out to be different from a conceptual point of view: ontologically speaking they cannot be the same. It is therefore possible that this interpretation changes over time while the source expression remains the same.

These terms can be categorized into those that are *in* the institutional reality created or modified by the legal rules, and those that are relevant to it but belong to brute reality from the point of view of the institution. These legally relevant terms may be institutional ones which belong to a separate institution whose existence is merely recognized by the institution the source of law applies to. Specific patterns involving the **constitutes** property separate these categories of terms in the knowledge representation.

Obviously multiple sources of law may be about the same institutional reality, and an explicit marker that this is the case is a desirable feature of legal institutions. One may also be of the opinion that legally relevant terms used in the source of law are the same as those used in other sources of law.

Since the same entity can be identified by multiple URI, it is possible to separate these interpretations from the codification of the terminology itself; All terms identified in a source of law expression are identified by a URI, for instance $n_1 : C_1$, that only identifies the term as it occurs in that source of law expression. All terms identified in a source of law are in the same namespace, and no other terms are in that namespace. The integration of terminologies from different source of law expressions is made explicit in the form of equivalence statements $n_1 : C_1 \equiv n_2 : C_2$.

The most obvious such scenario, and at the same time the most complicated one from a maintenance point of view, is localized change of a work: the set of terms in the new expression for a very large part overlap with the previous expression of the work. One might think of this as the shared set of terms at the work level, but the shared work level set only exists from a specific vantage point in time, or only once the source of law has become immutable after its repeal and the shared work level set – which can no longer change – has become irrelevant.

A pragmatic solution to reduce the number of such terms is to identify terms by the *expression* that created (first referred to) them in that work, and only to renew it if a new expression changes it by changing the set of OWL axioms that use it. This however creates a potential problem in relation to equivocation with terms from other works. It may happen, in rare cases, that institutional realities are split up.

Recently in Belgium, certain areas of legislation, including the existing works, have for instance be delegated to a lower level. What was once one federal institutional reality, has been split into three: the terms that remain on the federal level and retain a single meaning, and those that are now managed by the Flemish and Walloon community, respectively, and whose meanings now diverge.

This is however solved with the presumption that, with a change of author, the work is republished. This however requires a temporal fiction: for the correct application of *lex posterior* one has to presume that the community government is now the original author of all changes in the original timeline before the delegation happened. From a MetaLex point of view, it is certainly preferable to consider the delegation only a delegation of the power to change the source of law in the future, and to only relabel terms as soon as a divergence happens.

The legal rules and facts, together legal propositions, follow very similar rules. All propositions identified in a source of law expression are identified by a URI, for instance $n_1 : r_1$, that identifies the proposition as it occurs in that source of law expression. All propositions *created* in a source of law are in the same namespace, and no other propositions are in that namespace. A new URI is assigned only if the proposition changes in a new expression of a work.

Do note however that the source of law on the expression level *cites* other rules on the work level, while the legal rules we represent are necessarily identified by their *first representation* in an expression, and not on the work level. This creates a problem for legal rules *about* other legal rules: the logical representation changes not only when the rule is changed, but also if the rule it is about is changed. The maintenance problem this would create is easily addressed by using the notion of embodiment between expression and work in institutional reality. If we need to say that legal rule r for instance claims that some set of rules represented by expression level text fragments

that embody *work fragment* w apply to C , we can do that as follows:

$$\begin{aligned} \exists \text{representedBy}.\exists \text{metalex:embodies}.\{w\} \sqsubseteq \\ \forall \text{appliesTo}.(C \sqcap \exists \text{applicable}.\{r\}) \end{aligned}$$

The notion of embodiment applied in institutional reality is not very elegant, but it is the legislator itself that is the origin of confusion between institutional reality and the texts that prove its existence. It is this kind of observation that has led to the adoption of the *FRBR* distinction between works and expressions in MetaLex. The applicability patterns in section 5.2.3 are thus overly simplistic and have to be modified to accommodate embodiment, which could however only be introduced after the discussion of the source of law qua document and MetaLex.

Legal propositions are identified by OWL constants: it is absolutely essential to state that each proposition is different from all other known constants in contexts where there is a risk of unification of constants. Both legal propositions and terms, which are OWL concepts, are always denoted by URI in OWL by default. Logical sentences are not. Note that it is not customary in OWL DL to explicitly assign a full URI to OWL axioms. They do however always have one, derived from the *xml base* of the document they are stored in, which will by default function as namespace, and they can be explicitly reified to assign one. The logical sentences are not found in the source of law: they are created by the knowledge engineer. The logic of grouping their identifying URI into namespaces, and the sentences themselves into files, and doing version management on those files, perhaps by making a version number or publication date part of the namespace identifier, remains solely a knowledge engineering concern, limited only by the agreed upon update mechanism.

5.4.2. Knowledge Engineering Considerations

The knowledge engineer must keep in mind *why* he is making a change in an OWL file that represents a source of law: is it because institutional reality changed, or because his interpretation of the sources of law changed? In the latter case, the change can be made without further ado. Representation of a source of law expression consists of two interpretative phases in which mistakes can be made:

1. Identifying the entities in institutional reality represented or created (first represented) by the source; and
2. formulating the logical rules that describe the constraints on institutional reality and the interfaces by which institutional reality can be changed, as evidenced by the sources and the way they are used.

Case law has an important role in our increased understanding of law over time, and at the same time creates ambiguity because it functions both as a source of information on the meaning of a source of law, and as a source of law itself. It can change our interpretation of institutional reality, but it also changes institutional reality by representing new rules. This conclusion is inescapable: also in civil law

jurisdictions case law has the potential to change institutional reality. How and why will be explained in section 6.7.2.

A specific goal of this book is liberating knowledge engineers from the self-imposed straightjacket of trying to represent a legal rule by a logical rule. The logical rules that can be impacted by a change of the source are simply the ones that refer to the propositions and terms that were changed. In the case of legal rules, this use is always mediated through specific patterns of use of the **applicable** property.

Note that legal “facts”, represented as proposed in this book, are always rephrased to rules because they only have effect if applicable. propositions which only postulate the existence of something are simply the first reference to an entity in institutional reality, and create it merely by referring to it.

The knowledge representation captures the entities that exist in the institution, including their temporal features, and the entities themselves are linked to the sources as evidence of their existence through the inverse of the **refersTo** and **represents** properties.

The goal of a book like this one is not only to retain *descriptive fidelity* in representation, but also to increase *intercoder reliability*. The patterns of logical rules in this chapter and in chapter 6 in this sense function as a limited menu of options for representation that increases the chance that knowledge engineers will represent the same sentence in a source of law in a similar way.

A second instrument for increasing intercoder reliability is the use of core ontologies imported by the knowledge representation of the source of law, and to define terms in the source of law as subtypes of the terms in the core ontology. As pointed out in section 3.5.2, one should however be cautious in choosing a core ontology: many generic knowledge components are not suitable for this purpose, since they represent abstract theories that embody a specific *epistemic competence*.

The primary distinction between knowledge sources for the LKBS remains the distinction between descriptive ones, and autoepistemic ones. The descriptive part of the knowledge base, ordered from more reusable to less reusable from the perspective of representation of sources of law, consists of:

1. the core ontology; MetaLex and LKIF play this role in this book, although we have liberally deviated from LKIF;
2. the ontological constraints and requirements found in the sources of law, which introduce the relevant terms, subsumed by terms in the core ontology, which is the only thing imported by ontologies of sources of law;
3. a judicious set of subsumption or equivalency statements that connect terms describing institutional reality in the respective sources of law;
4. a judicious set of subsumption or equivalency statements that connect terms describing brute reality in the respective sources of law; and
5. a judicious set of subsumption or equivalency statements that connect terms in the source of law to relevant external theories – legal, abstract, commonsense, etc. – that are imported at this point.

The similar, bit shorter, list can be made for autoepistemic sets of rules:

1. the rules found in the sources of law, which appeal only to terms in the sources

- of law;
- 2. rules found in other sources of law, or customary in the involved legal system, that one cannot do without;
- 3. the rules of the relevant external theories – legal, abstract, commonsense, etc;
- 4. a judicious set of common sense rules that apply to brute reality in the respective sources of law, in particular those required by useful external theories (frame assumptions, etc.); and
- 5. LKBS-specific knowledge availability rules and constraints.

Although it is good strategy to separate the representation of individual sources of law from the integration of respective institutional and brute realities (as advocated in [42]), judicious sets of subsumption or equivalency statements, in line with current legal practice and case law, are obviously very valuable reusable knowledge components.

A lot of valuable advice on integrating ontologies, versioning ontologies, and pruning them to throw out unnecessary ballast for a specific KBS exists (cf. for instance [226, 173, 210]). One should however be cautious with the idea of *merging concepts* often found in this literature: merging concepts instead of declaring them equivalent in itself never gains any computational efficiency. The case is different when two ontologies follow different approaches to doing essentially the same thing, i.e. follow different core ontologies: in this case real gains can be made by merging.

5.4.3. Applicability and Knowledge Sources

The isomorphic approach to representing sources of law is simple: take a source of law and represent it. Making an LKBS that exhibits a certain epistemic competence, for instance to determine one's taxable income and the amount of income taxes due, or whether a certain plan for a shed is eligible for a permit, is a lot more complicated. In this case we want to know which (fragments of) sources of law are applicable to the problem; We want to represent only these sources of law, or if a complete repository for such representations exists, we want to know which ones to import.

Wouldn't it be convenient if there was some sort of metadata set for knowledge sources representing sources of law that helps us to find the relevant ones for our domain of law?

We don't have to look very far for a candidate solution: applicability rules do this thing for the law. The exact same problem is of course faced by anyone who tries to find the rules relevant to some predicament. Applicability rules are used to:

- 1. avoid having to repeat the same requirement a lot of times;
- 2. determine from when to when a rule can be applied, and from when to when the events must have happened to which it is applied;
- 3. demarcate the extent of one's claimed jurisdiction over people, territory, and substance;
- 4. restrict the applicability of rules made by a lower legislator, often prior of delegation of legislative competence to the lower legislator; and to
- 5. make the application of one rule conditional on the application of another rule, or in other words, to force a choice between two rules (cf. section 6.6).

All uses except perhaps the last one have a direct relevance to the problem of deciding whether a source of law is relevant to us, and it is especially convenient if we find them in the first section of the source of law. Applicability statements are however found within a source of law and in other sources of law *about* the one we are evaluating.

The suggested representation in section 5.2.3 makes both the carrier of the applicability statement and its target explicit, so it is not hard to find them if one keeps proper track of the identity of one's sources and legal rules. As noted in section 5.3.1 it is often the case that relevant metadata of one source of law is found in another one: one of the rationales for the design of MetaLex is to make this connection explicit.

Very often applicability rules do not apply to a specific rule, or the set of rules found in a chapter, or source of law. Mandate and delegation (cf. section 5.2.5) may link it to specific offices and legislative competences. For instance the provision *the agency for legal knowledge engineering creates additional guidelines for the representation of legal knowledge* creates the possibility of creating legally recognized rules for a newly invented agency (the agency for legal knowledge engineering), and restricts the applicability of those rules to the act of representing legal knowledge. It means in other words that the set of rules that is the result (**metalex:result**) of an act of the agency for legal knowledge engineering is restricted by the substance to which it can apply.

Applicability as a guide for clustering sources of law into domains of application works very well. The Dutch government search engine for legislation²⁴ is based on tracking legislative mandates, and is able to show for any formal act which lower regulations are based on it. There are also ways of indirectly doing the same, for instance grouping sources of law by formal author; the agencies and offices that compose a government are generally created to exercise specific public competences, often including legislative ones.

In section 5.2.4 it was noted that *legal fiction* and *also-applicability* have the unfortunate consequence of requiring us to add an extra condition to other logical rules that we prefer to be immutable. Besides that it was considered a form of fraud in law by Bentham. Of course one could defend the point of view that the consequences for the proposed logical formulas is simply evidence for the non-adequacy of this representation, but the fiction has wider implications for the problem of finding the rules that apply to one's predicament: it expands the context of applicability of other rules in counter-intuitive ways.

Adding non-indicative rules – the rules that restrict the valid models of the institution instead of the interface by which it is changed – should always have the effect of restricting the set of valid models. The fiction does the opposite: it expands the set of valid models. The correct way to do that is to *retract* rules, and replace them with less restrictive ones. The fiction can expand jurisdiction, expand the applicability of a source of law, and expand the applicability of a knowledge source representing a source of law.

The *ex tunc* modification of a source of law, which exists in countries which have a

²⁴<http://wetten.overheid.nl>

constitutional court that can annul legislation after it went into effect²⁵ has a similar effect of changing the past, and retroactive applicability also does potentially (cf. section 5.2). In this case – because these are predictably patterned fictions that occur a lot – they have been accommodated by three layers of timelines:

1. the interval in which the occurrences to which a legal rule can be applied must have happened;
2. the interval in which the action of applying the legal rule can happen; and
3. the interval in which a source of law has been known to exist (*ex tunc*).

These examples show that fiction patterns *can* be accommodated without accepting nonmonotonic reasoning in the institutional ontology, but we need to engineer a special solution for each such fiction pattern. Arbitrary use of fictions is very disagreeable from a knowledge management point of view.

²⁵The concept of *judicial review*, as contrasted to *legislative supremacy* or *parliamentary sovereignty*. In the EU many members appear to have such a possibility; Finland, the Netherlands, and the UK at least do not.

Chapter 6

The Normative Order

6.1. Introduction

The function of the institution, to create and formalize normative order, is different from its structures and rules of change. The function of artifacts is generally more important than how they work: there are for instance considerably more people who know what a VCR is good for and how it is operated, than people who know how it works internally.

The legislator believes that by shaping the institution in certain ways certain beneficial effects are brought about in brute reality. The rules of the institution directly or indirectly reflect preferences of the legislator with respect to the behaviour of other agents and assumptions of the legislator about the preferences and abilities of other agents; The relevant mechanisms were described in sections 4.5.2, 4.6, and 4.7.

When we interpret the source of law in terms of agent roles and scripts, of obligation, or of subjunctive betterness we are in effect taking into account the intentions and expectations of the legislator.

Legal knowledge representation does not merely reflect the institutional character of the rules, but often also tries to capture the intended normative order: creating it is the function of the institution. The most obvious way in which the legislator creates normative order is through normative rules: in legal knowledge engineering these are often represented in deontic logics.

Section 6.2 presents the representation of normative rules. The normative rule *qua* normative rule presents us with little difficulty, but also fails to address planning in compliance with the rules.

While chapter 4 only discussed the effect of rules on behaviour in isolation, and chapter 5 discussed logical form and applicability, this chapter also introduces and discusses several phenomena – normative conflict (see section 6.3 and 6.3.1) and contrary to duty norms (see section 6.4) – that frequently occur in sets of norms *when interpreted as a coherent normative order*.

When normative rule n is considered in isolation a classification into cases in $|\exists \text{allowedBy}.\{n\}|$ and cases in $|\exists \text{disallowedBy}.\{n\}|$ suffices to guide behaviour; Because it is relative to the individual rule n only no such thing as normative conflict arises. When normative rules are however considered as a coherent set used to order alternatives in a concrete decision problem, normative rules can expose a betterness relation that allows for gradations of goodness, as the example of the contrary-to-duty norm shows. This betterness relation moreover need not be unambiguous even in legislation from a single legislator, as the existence of normative conflict shows.

A coherent explanation of these phenomena has to appeal to either the account in terms of normative positions, or in terms of a betterness relation, and one has to take the position that certain sets of normative rules are formalizations of coherent ideas of an intended normative order. As pointed out in chapter 4, these accounts are abstract theories *fitted on* normative rules.

This betterness relation cannot be a relation between OWL DL classes, however, as suggested by the informal explanation in section 4.6, but must be a straightforward OWL property to be named \preceq . Section 6.5 discusses this OWL property as if it were a new modal logic-based logic of betterness, in keeping with the interpretation of OWL

DL as a labeled multimodal logic in section 3.4. Because the account would only apply to specific sets of normative rules, the \preceq property is subscripted \preceq_n to make this distinction. The account in terms of betterness *at least* applies to each individual normative rule, but has little added value if only applied to individual rules.

Note that the same OWL translation issue would apply for a classical deontic account in terms of O and P modal operators, which would normally be interpreted as $\forall.P$ and $\exists.P$ for some relevant property P . Because OWL DL lacks reflexivity, required for a classical deontic logic, this would be a dead end.

Section 6.6 discusses the more general problem of having to choose between legal rules; choice rules allow or disallow certain choices. Since the choice between legal rules is often motivated by the ascribed intended normative order, i.e. the normative conflict, this chapter is the right place to introduce the choice rule (see section 6.6.3), which is a special kind of applicability rule. This chapter does not however try to give an account of when and why choice rules are introduced in cases of normative conflict: the applicability of a choice rule does *not* depend on a recognition of normative conflict. Section 6.6.2 points out an example where such choice arises without any form of conflict.

Section 6.7 discusses the intended normative order from a number of perspectives: the legislator's, the court's, and those of addressees of the source of law and those involved in transaction with them, who base preferences and expectations on the prevailing normative order as understood by them.

The rules of the institution have an instrumental function for those that engage in social interactions governed by those norms that goes beyond their interpretation as constraints on behaviour, as explained in section 4.5.2 and section 4.8. In section 6.7.1 the interpretation of rules as found in the source of law as ingredients for action scripts is discussed.

Since the legislator is neither omnipotent nor omniscient in his abilities to bring about normative order in brute reality, his attempts to do so will inevitably fail occasionally.

Intentions have a habit of coming to the foreground when attempts and failures matter: sometimes even courts appeal to the intentions of the legislator to explain how legislation should be applied (cf. for instance [19]). Perceived failures of the legislator – commonly in the form of perceived normative conflict – creates the space for courts to get involved in creating normative order themselves, as section 6.7.2 explains.

The courts are the legitimate authors of choice rules.

Section 6.7.3 considers the values the legislator tries to promote through the intended normative order: these unsurprisingly also take the form of propositions of subjunctive betterness, the *norms of analysis*. Section 6.7.3 points out a weakness in the prevailing view of comparing legislation based on ontology integration; Comparative law in actuality compares intended and actual normative orders relative to selected social scripts (as noted in section 4.5.1), instead of comparing institutional ontologies. If knowledge engineers limit themselves to the institutional interpretation of legal rules, no meaningful comparison can be made.

6.1.1. Contrast and Subjunctive Betterness

The reading of obligation $O(\alpha \mid \beta)$ given in section 4.6 is *if β , it would be better if α* , as opposed to $\neg\alpha$.

If we take the account of obligation to represent either a single normative rule or a selected set of them we have to make a refinement: $O_i(\alpha \mid \beta)$ means *if β , it would be better if α* , as opposed to $\neg\alpha$, with reference to axis i , which is the standard used for comparison.

In section 6.2 we will distinguish the readings *if $\beta \sqcap \neg\alpha$, it would have been better if $\beta \sqcap \alpha$* , and *if $\beta \sqcap \alpha$, no $\beta \sqcap \neg\alpha$ would have been better*. The same reading will be applied to preferences and values in section 6.7.

The normative rule involves four ingredients: the case to which it is applicable, the case which is disallowed, the case which is allowed, and the axis of reference being the identifier of the norm itself. Similar structures are found in *comparison theory*, *psychological geometry*, and *measurement theory*.

Levinson's comparison theory (cf. [188]) deals with *rhetorical comparison* or *simile*. Generally, a rhetorical comparison compares a thing A (the *primum comparandum*) to a thing B (the *secundum comparatum*) on the basis of a quality C (the *tertium comparationis*): A is like B with respect to C , or A is as C as B , versus A is different from B , A is less C than B , etc. Just like in deontic reasoning, the axis of reference or tertium comparationis (in deontic reasoning the norm) is more often than not missing in similes.

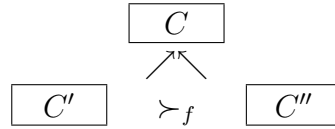
Kelly's *psychological geometry* (cf. [171]) is based on the supposition that we explain the world with theories built up from *personal constructs* that help us to find relevant contrasts (e.g. hostile vs. friendly, movable vs. immovable, etc.) relative to an *axis of reference*. According to Kelly "A person's [*mental*] processes slip into the grooves which are cut out by the mechanisms he adopts for realizing his objectives" (cf. [171], p. 49): to explain the situation is to make distinctions using these personal constructs. Kelly explicitly contrasts the better thing against its worse alternative(s). Measurement theory is similarly based on contrasts to define scales of measurement (cf. [242] for an overview and relevant references).

Repertory grids and associated methods (cf. [113]) for *knowledge acquisition* are based on psychological geometry: one for instance takes a topic concept, imagines three salient instances of it, and then tries to think of a distinguishing criterion present in two instances, but clearly absent in the third one, and then classifies these instances into two contrasting subconcepts along the exposed axis. To the subcategories the same procedure can be applied to produce a kind of taxonomy.

The criteria people come up with usually have an evaluative flavour: they distinguish a category that is "good" and a category that is "bad" for some purpose. The goodness of the category is obviously not a property of the category, but a judgment relative to an intended purpose of applying the construct to something. Constructs represent assessment criteria, and the taxonomies produced in this way reflect the way in which these criteria are prioritized: they are in effect handcrafted decision trees – perhaps even a kind of *scale* to assess things against – rather than true taxonomies, combining subsumption and betterness as organizing principles for knowledge

representation.

The result of such knowledge acquisition exercises are triangles between a general concept C and two subtypes C', C'' of it that 1) are disjoint ($C' \equiv \neg C''$), 2) completely partition the subsuming concept ($C \equiv C' \sqcup C''$), and 3) generally indicate that instances of C' are better than instances of C'' ($C' \succ_f C''$) for some purpose, or for the realization of some function f , graphically represented as:



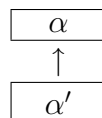
Interestingly, one of the mistakes commonly made by students who learn to build ontologies in a description logic like OWL DL is that they assume that subclasses are disjoint and fully partition the superclass by default, and are surprised that the reasoner fails to come to the expected conclusions (cf. [233]; also known from personal experience and that of colleagues).

Psychological geometry should certainly not be advanced as a method for designing ontologies: the reader will recognize that psychological geometry maps out the epistemological purposes of concepts and properties rather than a conceptualization of a domain detached from its immediate context of use. Psychological geometry is for instance a good method for building simple decision trees by hand, and structured comparison of sets of three instances is a good conversational tool for knowledge acquisition. It is also a suitable basis for a representation of subjunctive betterness.

In LKBS relevant contrasts also play a central role. Particularly interesting is for instance the notion of case factors favouring defendant or plaintiff in case law (cf. [18, 19] for a clear formulation, or section 6.7.2). A factor is not a plain fact: the facts of the case present in the description given in a verdict are present exactly *because* they favour one of the sides, and they are therefore factors. The justification of the case decision obviously reveals selective attention to those facts deemed to be (potentially) relevant for the decision, and they are relevant for the decision because they are relevant for the normative order the legal system intends.

Since repertory grids and decision trees are more or less recognizable conventions, and moreover seem to have a certain intuitive validity, this chapter, in particular section 6.2, will use Kelly's triangles liberally as a conceptualization of betterness. Note however that Kelly's triangles suggest a preference relation on propositions, and is therefore an epistemic account that requires reification of concepts. In OWL DL this reification is not allowed.

This chapter uses some notational conventions in addition to those introduced in chapter 3 for the use of Kelly's personal construct triangles. Because subsumption relations play a key role in this discussion, the following graphical convention for displaying a subsumption relation $\alpha' \sqsubseteq \alpha$ is occasionally used:



The set $|\phi|$ is the set of OWL individuals i such that $M, i \models \phi$. Another notational convention is used to make clear that a statement $|\beta \sqcap \alpha| \succ |\beta \sqcap \neg \alpha|$ is an ordering on α and $\neg \alpha$ within the context of β :

$$\beta : \boxed{\alpha} \succ \boxed{\neg \alpha}$$

An alternative notation is $|\beta \sqcap \alpha| \succ |\beta \sqcap \neg \alpha|$ or the following:

$$\boxed{\beta \sqcap \alpha} \succ \boxed{\beta \sqcap \neg \alpha}$$

Personal construct triangles are however only used for explanatory purposes.

6.2. Representation of Normative Rules

The normative rule is a legal rule that constrains assertions of the form **disallows**(n, a) and **allows**(n, i). The properties **allows** (inverse **allowedBy**) and **disallows** (inverse **disallowedBy**) represent a relation between legal rules and actions. Because of these domain restrictions, and the obvious disjointness of legal rule and action, the properties are irreflexive and asymmetric.

The two properties are disjoint, and subproperties of **appliesTo** (inverse **applicable**). If a rule allows something, it does not disallow it. If a rule allows or disallows something, it is also applicable to that thing.

The ontology fragments introduced here are largely part of the *LKIF* ontology (cf. [35]), although this chapter includes some new notions.

Normative rules are modeled by *constitutive* rules that make the normative rules *apply* to actions, and qualify them as allowed or disallowed. These rules are distinguished from other constitutive rules by their function: they separate the good from the bad, and the worse from the better, from the perspective of the legislator. Their function is to influence choice: the addressee of the rule is to choose allowed alternatives over disallowed ones. Normative rules are intended to have a *deontic choice* effect.

Constitutive rules normally constrain assertions of the form **constitutes**(b, i), where b is a brute thing and i an institutional one. One could think of the forms **disallows**(n, a) and **allows**(n, i) as alternative renderings of **constitutes**(i, v_n), **constitutes**(i, u_n), where v_n is a violation of n and u_n a use of n . This representation would however be needlessly verbose and complicated.

The representation of normative rules proposed here is based on a slightly different mapping from deontic operators to the ordered set $\{\textit{allowed}, \textit{silent}, \textit{disallowed}\}$ than the one in [265], but is otherwise inspired by it. Just like in the treatment of applicability given in section 5.2.2 I am however interested in representing explicitly *which* legal rule was violated.

The purpose of the representation is create the following mapping:

$$v(c) = \begin{cases} \textit{disallowed} & : O(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{disallowed} & : F(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{allowed} & : O(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{allowed} & : F(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{allowed} & : P(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : P(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{silent} & : \textit{otherwise} \end{cases}$$

The difference is mainly relevant in relation to so-called *compliance conflicts* (cf. section 6.3). The view that prohibitions and obligations give rise to a *strong permission*, contrary to the views of Kelsen in [172], Lindahl in [189], and Valente in [266], who believe that the obligation gives rise to a *weak permission*, will be defended in section 6.3.1.

The representation of normative rules models each legal rule by multiple constitutive ones. All rules state something *about* n instead of purporting to be *the* representation of n , taking into account the criticism expressed in section 2.2.2, and easily accommodate alternative representations of the things allowed or disallowed. This distinction between the legal rule and the logical propositions describing it was not explicitly made in [265].

Proposition 26. For each legal rule n interpreted as an obligation $O_n(\alpha \mid \beta)$ or prohibition $F_n(\neg\alpha \mid \beta)$, create three necessary conditions and three indicators (r_1, r_2, r_3):

$$\begin{aligned} \{n\} &\sqsubseteq \forall \textit{appliesTo}.\beta \\ \{n\} &\sqsubseteq \forall \textit{disallows}.\beta \sqcap \neg\alpha \\ \{n\} &\sqsubseteq \forall \textit{allows}.\beta \sqcap \alpha \\ (\text{DEFAULT } r_1 \text{ (known}.\beta)(\textit{free}.\exists \textit{applicable}.\{n\})} \\ &\quad (\textit{assume}.\exists \textit{applicable}.\{n\}) \\ (\text{DEFAULT } r_2 \text{ (known}.\beta \sqcap \neg\alpha)(\textit{free}.\exists \textit{disallowedBy}.\{n\})} \\ &\quad (\textit{assume}.\exists \textit{disallowedBy}.\{n\}) \\ (\text{DEFAULT } r_3 \text{ (known}.\beta \sqcap \alpha)(\textit{free}.\exists \textit{allowedBy}.\{n\})} \\ &\quad (\textit{assume}.\exists \textit{allowedBy}.\{n\}) \end{aligned}$$

Proposition 27. For each legal rule n interpreted as an obligation $P_n(\alpha \mid \beta)$, create two necessary conditions and two indicators (r_1, r_2):

$$\begin{aligned} \{n\} &\sqsubseteq \forall \textit{appliesTo}.\beta \\ \{n\} &\sqsubseteq \forall \textit{allows}.\beta \sqcap \alpha \\ (\text{DEFAULT } r_1 \text{ (known}.\beta)(\textit{free}.\exists \textit{applicable}.\{n\})} \\ &\quad (\textit{assume}.\exists \textit{applicable}.\{n\}) \\ (\text{DEFAULT } r_2 \text{ (known}.\beta \sqcap \alpha)(\textit{free}.\exists \textit{allowedBy}.\{n\})} \\ &\quad (\textit{assume}.\exists \textit{allowedBy}.\{n\}) \end{aligned}$$

The syntactical operations translate normative rules represented by the source of law, interpreted as obligatory and permissive rules, respectively. Just like the formalism in [265] these should not be interpreted as more general accounts of obligation and permission. All rules depend on only three ingredients: n, β, α . It is trivial to generate them syntactically from a user interface implementing just one representation that involves all three ingredients, for instance the classical $O_n(\alpha \mid \beta)$.

Recall from section 4.7 that normative statements in principle always address actions, although the intended category of actions may only be described in terms of a relevant result and not the way in which it was brought about. The formulation n : *in situation β you ought to do α* for instance becomes:

$$\begin{aligned}
& \{n\} \sqsubseteq \forall \text{appliesTo}.\exists \text{situation}.\beta \\
& \{n\} \sqsubseteq \forall \text{disallows}.\exists \text{situation}.\beta \sqcap \neg \alpha \\
& \{n\} \sqsubseteq \forall \text{allows}.\exists \text{situation}.\beta \sqcap \alpha \\
& (\text{DEFAULT } r_1 \text{ (known}.\exists \text{situation}.\beta)(\text{free}.\exists \text{applicable}.\{n\}) \\
& \quad (\text{assume}.\exists \text{applicable}.\{n\}) \\
& (\text{DEFAULT } r_2 \text{ (known}.\exists \text{situation}.\beta \sqcap \neg \alpha)(\text{free}.\exists \text{disallowedBy}.\{n\}) \\
& \quad (\text{assume}.\exists \text{disallowedBy}.\{n\}) \\
& (\text{DEFAULT } r_3 \text{ (known}.\exists \text{situation}.\beta \sqcap \alpha)(\text{free}.\exists \text{allowedBy}.\{n\}) \\
& \quad (\text{assume}.\exists \text{allowedBy}.\{n\})
\end{aligned}$$

The proposition $\exists \text{situation}.\beta$ should be interpreted as *doing something in situation β* .

Instead one may find other types of constraints on action, for instance a modification of the action, n : *when doing A you ought to do it in a A' manner*, or indication of a change, for instance n : *in situation S , S' should be initiated* or n : *situation S should be terminated*, that are all self-evident.

This representation is very similar in form to Valente's in [265], although [265] expressed the opinion that norms apply to situations.

Straightforward case assessment can be performed based on these constitutive rules alone. We can determine which known rules allow, and which known rules disallow, something s through a query $M_{\mathcal{DL}} \models \text{allows}(n, s)$, $M_{\mathcal{DL}} \models \text{disallows}(n, s)$, for all n in the abox. In [265] a criterium for recognizing conflict (informally discussed in section 6.3.1) between normative rules is supplied and a simple conflict resolution method is given. This method could in principle be applied here, which we will not do.

The approach presented here has the advantage over [265] that it is completely represented in the description logic: relevant subsumption relations between components of normative rules are automatically computed.

This would however require us to make a distinction not made so far if the conflict resolution method in [265] would be used: the norms subsume each other only with respect to their propositional content.

Subsumption between normative rules would be a bigger problem. Let n and m be two different norms. It is not sensible to think that $\{m\} \sqsubseteq \{n\}$: this would mean

that $m = n$. So there must be a generalized norm M and N , such that $\{n\} \sqsubseteq N$ and $\{m\} \sqsubseteq M$, and $M \sqsubseteq N$, and to these all these axioms about content apply, while specifics of $\{m\}$ and $\{n\}$ do not. M and N are the *norms qua qualification*. These must strictly speaking be different rules: surely N is an abstraction of $\{n\}$.

It is important to realize (as pointed out in section 3.4) that in OWL logical constants – in reality uniform resource identifiers – like n are not bijective. Given two norms n and m , it is permitted to state that $\{m\} \equiv \{n\}$ which will only succeed if the combination of both norms is a consistent norm. This line of exploration will however not be followed any further in this book.

The use of constitutive rules instead of institutional ones for normative rules is in first instance a principled one. Whether some state of affairs is allowed or disallowed can hardly be considered to be defining of the structures a legal institution can take on. The institution can certainly take on forms disallowed by some normative rule. It is not ontology, but an opinion on the relative desirability of the things qualified.

There are however certain logical relations between allowed and disallowed, *within the same axis of reference*, that are ontological in character: the same thing is not allowed and disallowed, and things allowed are better than things disallowed, relative to the same axis of reference.

This is a reason to place the allowed vs. disallowed distinction in its own institutional reality, just like Searle does by distinguishing between the institution of chess and the institution of competitive game playing to which winning and losing belongs (cf. [248]). One also cannot win and lose *the same game*.

There is also a practical precautionary reason to reject the interpretation of normative qualification rules as (potentially) institutional rules. Normative rules are clearly subject to applicability conditions and an arbitrary requirement that can function as exclusion grounds: inference to violation of a normative rule is therefore at least potentially defeasible without the existence of normative conflict.

The normative rules are constitutive, but the notion of *normative conflict* is *not* addressed as a standard default reasoning feature as is often the case in nonmonotonic deontic logics. The rules only state that some things are (dis)allowed by n ; No claim is made about a normative order as a whole (in keeping with Makinson's argument quoted in section 2.2.2), and normative rule n presumably never conflicts internally. Normative conflicts only give rise to defeasibility subject to certain assumptions.

The only type of logical inconsistencies to be expected without making assumptions about normative order involve contradictions between for instance a normative rule n : *blue cars are not allowed in this parking area* and an applicability rule *rule $\{n\}$ only applies to red cars* or some other requirement. The normative rule is subject to exceptions, but these are different in character than normative conflicts.

The method given in [265] did not address applicability conditions and exceptions at all. Instead these had to be expressed as more specific normative rules, forcing a normative conflict.

When we use a set of constitutive rules as proposed above, we do not account in any form for an interpretation in terms of betterness. The qualifications allowed and disallowed however have a logic of their own, because they are evaluative concepts.

As stated above, but also in section 6.1.1 about contrasts, we expect the set of

things that are allowed $|\exists \text{allowedBy}.\{n\}|$ and the set of things that are disallowed $|\exists \text{disallowedBy}.\{n\}|$ by a single normative rule to be disjoint. This property is inherent in the syntactical translation proposed.

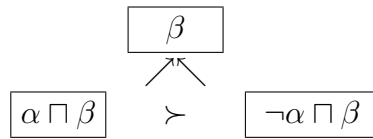
We moreover expect that things allowed are better than things disallowed, relative to the same axis of reference. This expectation, important for complying with the normative rules in planning applications is a lot harder to enforce in OWL DL. For now we write down this expectation as follows, where \succ_n is a relation *better relative to axis of reference n*:

$$|\exists \text{allowedBy}.\{n\}| \succ_n |\exists \text{disallowedBy}.\{n\}|$$

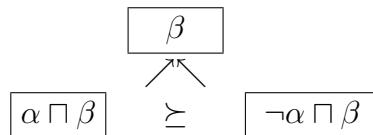
In planning we generally have to comply with multiple relevant rules. If we interpret these rules as *a set* of ordering constraints on alternative plans – in lieu of actions to which the rules apply – we have to deal with two important complications: firstly that the betterness relation is not antitransitive, which would be convenient, and – more importantly – that it is also not actually asymmetric.

Section 6.3 informally discusses normative conflict, a special property of some occurrences of symmetry in the betterness relation, and section 6.4 discusses contrary-to-duty norms, which are cases of apparent transitivity of the betterness relation.

In the following sections I will interpret a normative rule $O_n(\alpha \mid \beta)$ or $F(\neg\alpha \mid \beta)$ as:



The permissive normative rule cancels an obligation or prohibition, hence $P_n(\alpha \mid \beta)$ is interpreted as:



Note that it is the practical predicament of planning, and expectations about the rationality of the legislator, that suggests that certain *sets* of normative rules *should* represent a coherent betterness relation. The \prec and \preceq relation must in reality be a relation between individuals, and be restricted to for each individual normative rule n to domain $|\exists \text{disallowedBy}.\{n\}|$ and range $|\exists \text{allowedBy}.\{n\}|$.

Section 6.5 proposes a betterness relation that *can* be used in OWL DL on sets of norms. It does not allow for normative conflict, which is in my view not commensurable with accounts of constraints in planning in terms of obligation. The existence of normative conflict rather proves that the account in terms of obligation does not fit on a certain set of normative rules, in the sense discussed in section 2.3.1. The contrary-to-duty obligation on the other hand is a normal thing that should be accounted for. In section 6.5.1 normative conflict, and the circumstances in which it can be said to arise, is discussed in relation to the formalism proposed in section 6.5. The solution

for dealing with normative conflict will only be introduced in section 6.6.3, and this solution does bring back defeasibility through the back door through the choice rule.

Large parts of the following three sections were published in [44, 43, 35], although this chapter also contradicts some of the conclusions of these publications.

6.3. Normative Conflict

An instrumental concept in explaining conflicts between norms is that of *realizability*; A norm is *realized* if the state of affairs it allows or disallows is the case. There is a conflict between a pair of norms if they are not *jointly realizable*.

A clear logical formulation of normative conflict is not easy to give: very different conceptions of what constitutes a normative conflict are given in the relevant literature. This section gives an informal exposition of types known in literature.

A distinction is usually made between so-called conflicts of disaffirmation and compliance conflicts. Lindahl (cf. [189]) defines disaffirmation as follows: “a relation between two norms of different deontic mode, one being permissive and the other mandatory”.

A disaffirmation conflict is in our context a circularity between a permission and either an obligation or prohibition. This occurs if a state of affairs is simultaneously allowed and disallowed. The first such situation – a conflict of disaffirmation between $O(\alpha|\beta)$ and $P(\neg\alpha'|\beta')$ where $\alpha' \sqsubseteq \alpha$ and $\beta' \sqsubseteq \beta$ – can be represented as follows:

$$\begin{array}{ccc} \beta : & \boxed{\alpha} & \succ \boxed{\neg\alpha} \\ \uparrow & \uparrow & \downarrow \\ \beta' : & \boxed{\alpha'} & \preceq \boxed{\neg\alpha'} \end{array}$$

This is called a *disaffirmation of an imperative* (cf. [30]). Intuitively it is meant as an exception and takes precedence to the primary obligation, but this is not necessarily the case as we will see in section 6.6. Some simple examples:

(1) Parking is prohibited in this neighbourhood. (2) License holders are allowed to park in the designated bays.

(1) Walking on the lawn is prohibited. (2) Park maintenance department employees are permitted to mow the lawn.

The second such situation – a conflict of disaffirmation between $P(\alpha|\beta)$ and $O(\neg\alpha'|\beta')$ where $\alpha' \sqsubseteq \alpha$ and $\beta' \sqsubseteq \beta$ – can be represented as follows:

$$\begin{array}{ccc} \beta : & \boxed{\alpha} & \succeq \boxed{\neg\alpha} \\ \uparrow & \uparrow & \downarrow \\ \beta' : & \boxed{\alpha'} & \prec \boxed{\neg\alpha'} \end{array}$$

This is called a *disaffirmation of a permission*. Intuitively this case should be handled in the same way as the previous one, with the disaffirming norm taking precedence. A simple example:

1. Parking is allowed in the parking bays.

2. Parking in the parking bays is not allowed for cars wider than 1,85m.¹

Disaffirmation of permissions and imperatives is a very useful tool for legislators. It allows the legislator to create exceptions to rules in sources of law, without amending sentences with yet another sentence fragment. Since sentences in legislation are already too long as it is, according to many, this is certainly a useful thing. Disadvantage of doing this is that it makes it harder to apply rules, because the task of checking whether there are any such exceptions anywhere in the legal corpus is left as an exercise to the reader.

Another purpose of the disaffirmation is to influence the legislation of another legislator without actually touching its legislation. If legislator A is superior to legislator B, a disaffirmation in the legislation of A will effectively amend the legislation of B. This function was explained in general terms in relation to *delegation*. Because the amendment now ends up in a completely different part of the relevant legal corpus, this type of exception will be hard to find.

This does not mean that disaffirmation conflicts are always merely a design feature of legislation, or that the disaffirmation conflict is the same thing as an exception. There are several other types of disaffirmation conflicts that have no intuitive solution, or clear purpose. The first such situation – a conflict of disaffirmation between $O(\neg\alpha|\beta')$ and $P(\alpha'|\beta)$ where $\alpha' \sqsubseteq \alpha$ and $\beta' \sqsubseteq \beta$ – can be represented as follows:

$$\begin{array}{ccc} \beta : & \boxed{\alpha'} & \succsim \boxed{\neg\alpha'} \\ \uparrow & \downarrow & \uparrow \\ \beta' : & \boxed{\alpha} & \prec \boxed{\neg\alpha} \end{array}$$

The following two simple rules exemplify this case:

1. Using network facilities in the classrooms is prohibited.
2. Using WiFi in the university building is permitted.

It is not clear which of the two rules is an exception to the other one, granted that we share the belief that using WiFi is subsumed by using network facilities, and being in the classrooms is subsumed by being in the university building. The second such situation is created by reversing the deontic modalities in the example:

1. Using network facilities in the classrooms is permitted.
2. Using WiFi in the university building is prohibited.

This is a conflict of disaffirmation between $O(\neg\alpha'|\beta)$ and $P(\alpha|\beta')$ – where $\alpha' \sqsubseteq \alpha$ and $\beta' \sqsubseteq \beta$ – and it can be represented as follows:

$$\begin{array}{ccc} \beta : & \boxed{\neg\alpha'} & \succsim \boxed{\alpha'} \\ \uparrow & \uparrow & \downarrow \\ \beta' : & \boxed{\neg\alpha} & \preceq \boxed{\alpha} \end{array}$$

These situations are certainly conflicts from a *joint realization* point of view, and would certainly have been recognized as such by Lindahl (viz [189]) or Hill (viz. [147]). Arguably they belong to the Hill's *intersection conflicts* (cf. [105, 147]).

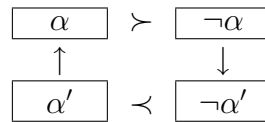
¹Inspired by the SUV amendment of 20 October 2004 of the town council of Nijmegen.

These cases defy Valente's formalization of normative conflicts (cf. [266]), which would not be able to determine which one is an exception to the other one. Any solution would rely on the distinction between settled context and that which is obliged that is typical for an ought-to-do perspective on norms (cf. section 4.7.1).

A special case is the *explicit disaffirmation*: the permission and the imperative disaffirm each other directly. This obviously happens if and only if $\alpha' \equiv \alpha$ and $\beta' \equiv \beta$. All four of the previous characterizations are applicable to this pair of norms. Intuitively this does not make sense at all. It is however in principle valid, provided that the two norms have been enacted by different legislators that are apparently disputing jurisdiction.

The compliance conflict is logically less transparent. The compliance conflict is defined by Lindahl as a relation between two mandatory norms, both of which are individually realizable, but not jointly. Simply put, the compliance conflict gives you a choice where none of the options is allowed, or all of them are obliged.

This situation – a *compliance conflict* between $O(\alpha|\top)$ and $O(\neg\alpha'|\top)$ where $\alpha' \sqsubseteq \alpha$ – can be represented as follows:



This essentially makes it a special case of disaffirmation. True compliance conflicts appear to be rare. Consider the following interesting real-world example of something which appears to be a compliance conflict: The Amsterdam police at some point ordered nightclub owners to keep emergency exits locked to keep drugs out, while the fire department ordered the same nightclub owners to keep them unlocked to allow for escape in case of disaster. Newspapers considered this a compliance conflict, and this one does indeed invite a logical representation that would make it one.

Whether this is a compliance conflict in the sense of “all alternatives are disallowed” however depends on how we conceptualize the alternatives of the nightclub owners. Strictly speaking, the nightclub owners did have the possibility of admitting much less customers, so that no emergency exits are required, or even to stop running a nightclub. Surely the correct reading of the obligation is that one should not initiate the situation of hosting a large number of customers and having one's emergency exits locked, or unlocked respectively. It only appears to be a case of “all options are disallowed” if one considers certain propositions to be part of a settled context, i.e. one does not have to consider alternatives that involve not admitting the same number of customers in one's nightclub as one did before. But where does this allowance for not considering that alternative come from? This kind of argument would not fly in many other contexts.

The treatment of compliance conflicts in automated reasoning systems as logical inconsistency is not really adequate, because the judgment call will depend on the inevitability, legality, and desirability of the settled context, and whether the agent appears to have intentionally or recklessly created a situation of perceived compliance conflict. Many perceived compliance conflicts are not compliance conflicts in a logical

sense, but only become one in a specific decision making setting in a settled context. This point of view will be developed further in section 6.3.1.

The conflicts or circularities between norms discussed in this section so far do not depend on the particulars of the case at hand in this formulation. The subsumption relationship between two norms does not come into existence in the context of a specific case. The simple fact that the same behaviour is permitted by one norm and prohibited by another, does not in any way imply a conflict. Translating norms to subjunctive betterness, and therefore conflicts to circularities, makes this clear, but the idea that a conflict arises between two norms with respect to a specific case is widespread.

In [266], Valente proposed a computation of the *prime implicant* (cf. the definition of prime implicant in section 3.3 or [215]) of the conflict. While Valente was right to observe that not any pair of norms of different deontic mode applying to the same situation point to conflict, it is unnecessary to involve an actual situation to be assessed in the computation, as observed in [286]. There must be in fact a straightforward subsumption relation between (relevant components of) the rules.

This means that to check for conflicts in a normative system, one does not need a batch of cases describing situations. Conflicts can be found – and usually resolved – using the norms alone. This cuts down dramatically on the amount of computation required for the solution in [266], because subsumption can be calculated offline and the results of it stored in a database for use by LKBS. Also, this point of view, which requires some more explanation, will be developed further in section 6.3.1.

As I have argued in [30], it is also extremely useful to recognize *affirmation of an imperative* and *affirmation of a permission* as a tool for structuring legislation. These closely related concepts are also based on subsumption, but do not involve a circularity or conflict. An affirmation of an imperative assumes two norms of the same mandatory deontic mode in a subsumptive relationship, and affirmation of a permission involves two norms of permissive mode. Affirmation as a logical instrument plays a role in the so-called *ne bis in idem* principle explained in section 6.6.

Example 12. An imperative for designing ships for instance states that equipment of any kind on ships should be designed “in such a way as to minimize potential damage of mishandling”. This norm is also *affirmed* in many places by other, more detailed, imperatives relating to specific bad user interfaces of specific types of equipment. It is for instance not allowed to use a single piping system with faucets to transport both water for the fire pumps and fuel oil for the propulsion system. Mistakes with faucets are easily made in a panic. It is likely the more general norm has been added because new ways of mishandling equipment are invented regularly as technology progresses, and designers are admonished to think of these scenarios before the ship is built. (example from [30]).

The previous characterizations of conflict are adequate when we consider only pairs of normative statements from law, but there are arguably other situations which are also conceptualized as a conflict that are not covered by this characterization. In some cases a pair of normative statements is realizable, but only in a way that is not satisfactory with respect to the intentions of the legislator. Perhaps we can tentatively

refer to this as the *conflict with legislator's intentions* scenario. One such scenario, constructed by Elhag et al. (viz. [105]), involves two permissions:

There seem to be other types of conflict as that between the permission for A to live in a certain house and a permission for B to destroy that same house. These conflicts need our attention and have to be embodied in a theory on normative conflicts.

It is trivial to think of similar situations involving two different agents and two norms. This situation is actually jointly realizable (to the detriment of A), and neither of the agents is confronted by a circularity in making his choice. In addition, both agents are free to act or not to act on the permission. As such, it is not really a conflict between the two involved norms, but everyone would agree that the resulting situation is awkward if both agents act on the permissions.

Another situation of perceived conflict arises if the legislator issues a norm that is *impossible to realize* or its realization does not depend on any agent choice. The legislator cannot order that dice should always land on six eyes, it cannot prohibit volcanoes from exploding, it cannot amend the law of gravity, it cannot prohibit fatal accidents, etc. If you ought to do something, you must have the ability to do that thing.

A last oddity is the lack of ability to apply a normative rule, of recognizing that it is violated or complied with. In the Netherlands there is a crime in the books, grievous blasphemy², that passed into desuetude when a court asked to apply it argued, with reference to the constitutional freedom of religion, that it found it impossible to decide what attributes are to be considered grievous when applied to a deity if the defendant denies having had the intention to grieve a deity. This is the more general situation of an indicative condition that never arises in brute reality, since blasphemers generally speaking do not have the intention to grieve the deity itself if the intention was to grieve in the first place. The rule was however never removed from the corpus since it is deemed to still have a function as moral guidance.

6.3.1. Compliance and Brute Reality

The traditional deontic axiom that states that everything that is obliged is allowed has always been a point of contention in legal theory and legal knowledge engineering. Kelsen in [172], Lindahl in [189], and Valente in [266] have stressed that obligations and prohibitions, besides classifying certain behaviours as disallowed, only render a *weak permission* to do the opposite of what is disallowed. In this view only the explicit *strong permission* classifies behaviours as allowed. Hence the following value function v mapping from descriptions to the ordered set $\{allowed, silent, disallowed\}$ (compare to the first table in section 6.2):

²Wetboek van Strafrecht, art. 147

$$v(c) = \begin{cases} \textit{disallowed} & : O(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{disallowed} & : F(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : O(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : F(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{allowed} & : P(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \alpha \mid \\ \textit{silent} & : P(\alpha \mid \beta) \wedge c \in \mid \beta \sqcap \neg\alpha \mid \\ \textit{silent} & : \textit{otherwise} \end{cases}$$

What this comes down to is that if you approach a T-junction and you see a sign saying that turning left is disallowed, then you have only a weak permission to turn around or turn right. Kelsen, Lindahl, and Valente do however recognize the existence of *compliance conflicts* (cf. section 6.3). If turning right or around is also prohibited, there is a conflict. These are however not considered conflicts between the implicit classification of the opposite of going left as allowed, and the classification of going right or turning around as disallowed, but as cases of joint unrealizability: two opposite behaviours have the same classification of disallowed. Valente's theory uses this notion of opposite to define the compliance conflict. There is no disagreement that one obligation may be used as a valid argument to violate another obligation.

The problem here is first and foremost a difference in conceptualization of conflicts, and secondly of what is the *opposite* of a behaviour.

If normative statements are not conditionalized, i.e. a statement $F\alpha$ is allowed to occur, the notion of a strong permission from a prohibition for instance is bizarre: the permission of the opposite $P\neg\alpha$ would cover almost everything. A prohibition stating that one is not allowed to walk on the grass surely doesn't allow walking anywhere else? The rejection of the implied permission because of the open-ended nature of the logical complement of the behaviour that is disallowed is however purely an esthetic one: the notion of "opposite behaviours" with the same classification creates exactly the same logical problem, but less obviously. This is however a red herring.

The fundamental difference between the accounts given in section 6.3 and [266] is the operationalization of conflict. The operationalization of section 6.3 is based on a logical conflict between two norms, applied to the same behaviour, regardless of the case at hand. The compliance conflict and the conflict of disaffirmation are really two faces of the same thing.

Valente's operationalization is far more general, and would indeed give bizarre results if obligations and prohibitions yielded strong permissions. In Valente's operationalization, the two involved norms do not have to stand in a subsumption relation to each other. In Valente's system, norms are applied not to a behaviour but to the *description of the case*: the thing classified by the norm is the *prime implicant* (cf. section 3.3) of the *generic case* description in the norm. This case description may describe (the results of) two or more different behaviours. Moreover the notion of prime implicant would only apply to the *description* of a behaviour, while the norms apply to the behaviour, at least in the view developed here.

What is required is that there is such a prime implicant case c which is classified as disallowed by some norm, while its negation is also a prime implicant case of another norm that classifies it as disallowed. Alternatively, for conflicts of disaffirmation,

required is that there is a prime implicant case c of two norms, one yielding disallowed and the other allowed. These two norms do not however have to be logically related.

The restriction to prime implicants obviously is obviously necessary to weed out false positives like “today I went to vote (which is strongly permitted) and then I shot everyone at the voting station (which is prohibited)”. But there is also a more profound truth found here. The question is whether the cases that would be covered by the prime implicant criterion, but do not involve a subsumption relation between normative statements, are true conflicts. This book takes the point of view that this is *not* the case, but there is a big problem with the restriction of irreconcilable *conflict* to subsumption only. The prime implicant criterion does pick up a category of cases that will sometimes be considered conflicts, but it also picks up false positives.

Clearly, the criterion “more specific” is wider than logical subsumption. See for instance also [168] for the same observation relating to prioritized defaults and the specificity criterion: intuitively one may want to consider a preference for *a film directed by Almodavar* as more specific than a dislike of *a Spanish film*, even though 1) they clearly aren’t in a subsumption relation to each other and 2) the whole notion of *a Spanish film* becomes underspecified when we consider Spanish directors, producers, cast, script, locale, etc. More specific in this sense is perhaps better understood as a refined level of granularity (as intended in section 2.3.1): there is a level of granularity at which the two norms would appear to be in conflict. This expanded notion of specificity is not reconcilable with the interpretation of obligations and prohibitions as overly general strong permissions, when considered in combination with conflict resolution strategies like *lex superior* and *lex posterior*³.

It is essential to note that the fact that two legal rules are in “irreconcilable conflict” cannot possibly be inferred from the rules themselves. Assuming that institutional rules and requirements cannot participate in a conflict, the things that any pair of indicative rules apply to only stand in a subsumption relation to each other relative to a specific *theory of brute reality*. This theory of brute reality may be an *ontology* of brute reality, in which case there is no problem with a translation of specificity to logical subsumption, but the involved situation recognition will in many cases also involve ampliative reasoning, abstraction, etc. The framework sketched in chapter 3 does not account for subsumption between autoepistemic rules, and it is not clear what such a subsumption theory would look like when applied to the problem of detecting irreconcilable conflicts between legal rules. An approach to default subsumption has been sketched by [274], but adopting this very broad approach would seem to generate false positives.

The legislator does not supply a theory of brute reality with his legal rules, and therefore cannot possibly foresee what combinations of normative rules will at some point in time be construed as being in “irreconcilable conflict”. The nightclub example in section 6.3 shows the problem. There is no logical justification for considering this a compliance conflict according to [172, 189, 266] or this book, unless one models the involved situation so that they meet the logical formulation of the conflict, which just hides the problem.

³Confer sections 5.1 and 6.6. Conflicts solved by *lex specialis* will generally not lead to problems since the overly general permission will be defeated.

The question is why one would take the trouble to reformulate the rules so that they appear to be in conflict. Is the compliance conflict just one logical principle, akin to the conflict of disaffirmation, or is it something else? The compliance conflict in a classical logical sense appears to be a very rare event, or even a non-event. But there is another soft principle at work which can also be described under the heading of compliance conflict:

Proposition 28 (Compliance Dilemma). In any decision problem, there must be a *known* option or set of options that is the allowed choice in that decision problem.

The selection of the appropriate menu of alternatives is of great importance. Given a set of specific options, the rules may impose a number of ordering constraints on them. Circularities may arise, but do not in themselves constitute a conflict. A problem arises if all conceivable alternatives are disallowed, or strictly worse than inconceivable, fictional alternatives.

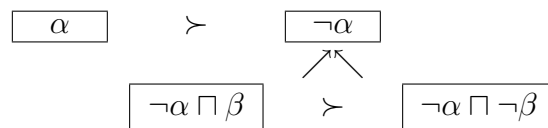
It is not good enough to claim – as a matter of ontology – that “there exists a better alternative that is allowed”. This better alternative must also be imaginable; There is a *burden of proof* on the prosecutor to supply such an option. If it isn’t, courts may latch on to a conspicuous circularity and use the principles that they normally use to resolve normative conflicts to construct a theory that allows them to conclude that no violation has taken place. This theory construction in the simple case involves *strengthening the antecedent* (cf. section 6.5.1) of the more generally applicable rule involved in the circularity.

It however makes a difference which decision problem we are considering. If proposition α is true in all imaginable alternatives, α is part of the *settled context* in that decision problem. To make a case that α is a violation one has to reconceptualize the decision problem, for instance by placing it back in time, when α was still preventable. Obligations that only apply in situations where the settled context already includes a violation are called *contrary-to-duty obligations*.

6.4. Contrary to Duty Obligations

The *contrary-to-duty* (CTD) obligation arises in a sub-ideal situation, brought about by the violation of a primary obligation. This is not allowed in standard deontic logic because it is based on a modal distinction between the actual world and the ideal one, and treats normative sentences as constraints on the ideal world. Treatment of obligations as a proposed betterness ordering of types of situations orders these worlds.

The relation between the primary obligation and *CTD-obligation* can be represented according to the following schema:



The CTD-obligation in the bottom row only distinguishes cases in which $\neg\alpha$ is already the case, in both allowed and disallowed situations.

It is very important to realize that the relation between the primary obligation and the CTD-obligation is not one of compliance or disaffirmation conflict. As long as you don't get into the sub-ideal situation, the CTD-obligation is silent – neither complied with nor violated. The CTD-obligation also does not imply a permission to enter into the sub-ideal situation.

The Chisholm paradox (viz. [77]) is an instructive example of how to use subjunctive betterness statements to order actions and to analyze complex sets of logical relations between primary obligations and CTD-obligations. The Chisholm set consists of the following norms:

1. $O(\alpha|\top)$
2. $O(\beta|\alpha)$
3. $O(\neg\beta|\neg\alpha)$

The logical paradox arises in standard deontic logic when α is the case. α reads “a man goes to the assistance of his neighbours”, and β reads “the man tells his neighbours that he will come”. Opinions are divided (viz. [265]) over whether this paradox already exists in natural language, or only in the logics that give rise to it. If one reads norms as statements of subjunctive betterness, one takes the position that the paradox merely disqualifies certain logical representations of the Chisholm situation.

The Chisholm set translates to the following set of preferences:

1. $|\alpha| \succ |\neg\alpha|$
2. $|\beta \sqcap \alpha| \succ |\neg\beta \sqcap \alpha|$
3. $|\neg\beta \sqcap \neg\alpha| \succ |\beta \sqcap \neg\alpha|$

The first preference expresses a preference for both $|\beta \sqcap \alpha|$ and $|\neg\beta \sqcap \alpha|$ over $|\beta \sqcap \neg\alpha|$ and $|\neg\beta \sqcap \neg\alpha|$. As for instance observed in (cf. [269]) a good formalization of ordering alternatives enforces the following ordering constraints:

$$\boxed{\alpha \sqcap \beta} \succ \boxed{\alpha \sqcap \neg\beta} \succ \boxed{\neg\alpha \sqcap \neg\beta} \succ \boxed{\neg\alpha \sqcap \beta}$$

An obligation is not cancelled by adding a second obligation telling us what to do if the first one is violated. If you tell your neighbours that you will come to their assistance, and don't do it, then you have violated two obligations. As a general rule, CTD-obligations do not require a choice between the involved norms. Straightforward deontic logics based on an accessibility relation interpreted as *ideality* often have a problem with this construct: such logics do not order the alternatives, but pick the ideal or, in some logics, minimally subideal one.

Only because the Chisholm situation involves a third norm ordering the two situations not ordered by the first two, we are able to construct a complete ordering. This is not generally the case with CTD-obligations.

Time adds another interesting twist to the Chisholm situation. Note that the transitions that can be effected by an agent's actions are only those from $|\neg\alpha \sqcap \neg\beta|$ to either $|\alpha \sqcap \neg\beta|$ or $|\beta \sqcap \neg\alpha|$, and from $|\beta \sqcap \neg\alpha|$ to $|\alpha \sqcap \beta|$. The choice for β can only be realized *before* the choice for α .

This makes a temporal reading, where norms only become visible at certain time points and norm 2 only becomes *visible* because norm 1 was violated, impossible (cf. [269]). This shows that the Chisholm example, although a bit contrived, has been chosen very well. A more straightforward and intuitive temporal example, by Prakken and Sergot in [232], admonishes one to keep one's promises, and to apologize if one doesn't.

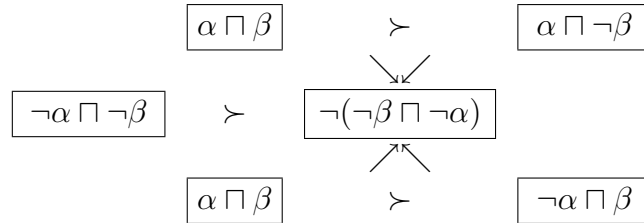
The Reykjavik set (cf. [78]) requires a partial ordering of the involved types of action. The Reykjavik set consists of the following norms:

1. $O(\neg\alpha \sqcap \neg\beta | \top)$
2. $O(\beta | \alpha)$
3. $O(\alpha | \beta)$

α reads “tell Reagan the secret”, and β reads “tell Gorbachov the secret”. This one is slightly more complicated to translate to preferences. The Reykjavik set translates to the following set of preferences:

1. $| \neg\alpha \sqcap \neg\beta | \succ | \neg(\neg\beta \sqcap \neg\alpha) |$
2. $| \beta \sqcap \alpha | \succ | \neg\beta \sqcap \alpha |$
3. $| \beta \sqcap \alpha | \succ | \beta \sqcap \neg\alpha |$

This shows us it is possible that there are more than one CTD-obligations pertaining to one primary obligation. The relation between this primary obligation and the two CTD-obligations can be represented in the following schema:



If CTD-sets get even more complex, it is obviously no longer possible to represent them graphically in this way.

The concept $| \beta \sqcap \alpha |$ is less cumbersome than $| \neg(\neg\beta \sqcap \neg\alpha) |$, but for the sake of making clear where it comes from the statement is not rewritten. The first preference clearly expresses a preference for $| \neg\beta \sqcap \neg\alpha |$ over $| \beta \sqcap \alpha |$, $| \neg\beta \sqcap \alpha |$, and $| \beta \sqcap \neg\alpha |$. The imposed ordering is a partial one:

$$\boxed{\neg\alpha \sqcap \neg\beta} \succ \boxed{\alpha \sqcap \beta} \succ \boxed{\neg\alpha \sqcap \beta} \quad \boxed{\alpha \sqcap \neg\beta}$$

The last two situations are left to personal preference for either Reagan or Gorbachov. Note that the transitions that can be effected by an agent are those from $| \neg\alpha \sqcap \neg\beta |$ to either $| \alpha \sqcap \neg\beta |$ or $| \beta \sqcap \neg\alpha |$, and from those to $| \alpha \sqcap \beta |$. There is no ordering constraint on choices between α and β .

Although temporal readings of CTD structures are ruled out, one might still speculate that the primary obligation and the contrary-to-duty obligation involve a substantially different decision point. In all examples given the decision points can in

principle be separated. The decision to tell or apologize is substantially different from the decision to come or break a promise. This intuition may be dispelled by Forrester's situation (viz. [111]), in which one is admonished to kill gently, if one kills. Killing gently is directly subsumed by killing, so it is not possible to kill gently while not killing. α and β are not at all independently realizable in this case.

The CTD situation should not be considered merely theoretical. Real legislation does actually create many situations reminiscent of the Chisholm and Forrester situation. Contrary-to-duty imperatives are quite common, although they are usually far more complex to analyze in legislation. All norms regulating punishment by the legal system, and contractual remedial and reparational obligations and liabilities arising from contract violations follow this general pattern.

The CTD situation is merely a recognition in legal theory of the fact that betterness relations are transitive. Since classical deontic logics are based on a distinction between what is the case and what would be *ideally* the case, the CTD situation forms a problem. Many modern deontic logics have addressed this issue; Some do so by explicitly modeling a preference relation.

6.4.1. Duty Conditional on Violation

There is a very different way of handling (some kinds of) CTD-obligations that should be mentioned. Instead of viewing a CTD-obligation as an obligation arising from the sub-ideal situation in which one norm has already been violated, one could see the CTD-obligation arising from the violation itself. Consider the following possible formulations of norms about apologizing when breaking promises in natural language:

1. One ought not to break a promise.
2. One may break a promise if one apologizes.
3. One ought to break promises and apologize.
4. If one breaks a promise, then one ought to apologize.
5. If item 1 is violated, then one ought to apologize.

Item 2 is clearly a permission that is an disaffirmation of item 1. Item 3 creates a compliance conflict with item 1. Item 4 is a CTD-obligation and at first glance expresses the same thing as item 5, but still works somewhat differently in a legal system.

Item 4 is only active as a CTD-obligation if item 1 exists. If the text of item 1 changes, then the preference expressed in item 4 changes as well. Governatori et al. (cf. [123]) solve this by the introduction of an operator $O(\alpha) \otimes O(\beta)$, which is read as " $O(\beta)$ is the reparation of the violation of $O(\alpha)$ ". This makes the obligation $O(\beta)$ inapplicable to any situations in which the opposite of α is the case, but $O(\alpha)$ is nevertheless not violated. The operator is therefore not suitable for replacement of a more general way of dealing with CTD-obligations. Governatori et al. may however have a valid argument that some CTD-obligations do actually require the violation. Arguably this is the case in the following set:

1. One ought not to break a promise.
2. One may break a promise if one pays a \$100 fee to the government.

3. If item 1 is violated, then one ought to apologize.

The question is whether item 1 is still violated if one pays the fee. Does one still need to apologize? We have a problem if that is not the case: we need to distinguish between CTD-obligations that are only applicable if the primary obligation is, and CTD-obligations that are applicable regardless of the primary obligation. This is a case of confluence of a CTD-relation on the logical level, and an applicability condition, as described in section 5.2.3.

Apparently the primary obligation does not need to be applicable if item 3 would have been formulated as “If the conditions of item 1 are met, then one ought to apologize.”, but there is no guarantee that it will be read this way.

This type of *duty conditional on violation* is for instance quite frequently encountered in contracts.

That the CTD-like duty conditional on violation n_3 only applies to violation of some other rule n_1 can be represented as:

$$\{n_3\} \sqsubseteq \forall \text{appliesTo}.(\exists \text{disallowedBy}.\{n_1\})$$

This necessary condition on n_3 simply that n_3 only applies to violations of n_1 . Contrary to the previous phenomena this pattern does not appeal to a perceived deontic choice effect.

6.5. Betterness in OWL DL

The contrary-to-duty situation is clearly a normal feature of legislation, and more importantly also a normal feature of decision problems, which indicates that \prec and \preceq are transitive. One would expect this of such relations. It presents a problem for simple modal deontic logics, which equate violation with inconsistency, and cannot deal with situations involving violations. Inability to deal with contrary-to-duty situations simply disqualifies certain logics.

The normative conflict is a different issue. It is a fundamental problem with the applicability of the obligation terminology on decision problems. In the presence of normative conflict, one can no longer be sure that one can strengthen the antecedent β of weaken the consequent α of an obligation $O(\alpha \mid \beta)$ (cf. for instance [268] for a discussion). There are of course many different ways to complicate deontic logic to accommodate normative conflict, but it is simply the concept of obligation, as understood in deontic logic, which no longer fits well on the decision problem. Interestingly, since permission is *intended* to cancel obligations, which it doesn't do in standard deontic logic, we can also conclude that the way standard deontic logic models permission makes little sense, as has been observed by many before.

Trying to model a simple deontic logic in OWL DL is therefore not very helpful. Even for a simple deontic logic OWL DL would moreover create problems, since OWL DL lacks the ability to enforce reflexivity on relations.

There are generally two routes one can take to address these issues:

1. Betterness (or preference) instead of ideality is a better fit for contrary-to-duty situations, and moreover connects better to decision problems like planning;

2. Defeasibility is an obvious explanation of the relations between obligation and permission, and of compliance conflict.

In this book I recommend both, but in different contexts. From an institutional perspective, normative rules are clearly defeasible rules as already indicated. Normative conflict is one of the reasons why normative rules are considered defeasible, dependent on an equally defeasible assumption about the pair of normative rules. Section 6.6.3 will discuss applicability rules that make the applicability of one normative rule dependent on the applicability of another one.

Obligation, just like preference, in first instance relates to decision problems. It is an account of choice between alternative plans. It is for this purpose that I introduce a property \succeq to model a relation between plans or actions. The notion of an *actual situation*, as distinguished with ideal ones, will however play no role. *LKIF* (see appendix to [35]) includes the same property.

Each normative rule posits its own unique betterness relationship, which can thought of as a norm in the proper contexts: a property \succeq_n is inferred from $\{n\}$. There is an expectation that these relationships can be combined into a single one for sets of rules in some way, resulting in an integrated theory of what the legislator wants us to do. The expectation in a standard deontic logic is very strong: a single \succeq can model each set of norms, and is besides that antitransitive. I do not attempt in this section to give a systematic account of these expectations.

Proposition 29. For each legal rule n interpreted as an obligation $O_n(\alpha \mid \beta)$ or prohibition $F_n(\neg\alpha \mid \beta)$, create the following terminological axioms:

$$\begin{aligned} \exists \text{allowedBy}.\{n\} &\sqsubseteq \neg \exists \preceq_n . \exists \text{disallowedBy}.\{n\} \\ \exists \text{disallowedBy}.\{n\} &\sqsubseteq \exists \preceq_n . \exists \text{allowedBy}.\{n\} \end{aligned}$$

Proposition 30. For each legal rule n interpreted as an permission $P_n(\alpha \mid \beta)$, create the following terminological axiom:

$$\exists \text{applicable}.\{n\} \sqcap \neg \exists \text{allowedBy}.\{n\} \sqsubseteq \exists \preceq_n . \exists \text{allowedBy}.\{n\}$$

Because we represent each normative rule by both an indicator and a requirement, $\exists \text{allowedBy}.\{n\} \equiv \alpha \sqcap \beta$ and $\exists \text{disallowedBy}.\{n\} \equiv \neg\alpha \sqcap \beta$ are “almost” true, barring applicability conditions that defeat the indicative rules. For discussion purposes we will assume that applicability of n already settled. Note however that continuous application of indicative rules will be necessary.

$O_n(\alpha \mid \beta)$ – barring applicability restrictions on n – therefore means:

$$\begin{aligned} \alpha \sqcap \beta &\sqsubseteq \forall \preceq_n \alpha \sqcup \neg\beta \\ \neg\alpha \sqcap \beta &\sqsubseteq \exists \preceq_n \alpha \sqcap \beta \end{aligned}$$

If the obligation is for instance *when driving a vehicle one should keep to the right of the road*, the interpretation in terms of subjunctive betterness is:

When driving a vehicle and keeping to the right of the road, all equal or better alternatives would have involved not driving a vehicle or keeping to the right of the road. When driving a vehicle and not keeping to the right of the road, an equal or better alternative would be driving a vehicle and keeping to the right of the road.

The combination of both statements is only satisfied by considering alternatives involving not keeping to the right while driving strictly worse than those that do involve keeping to the right, or by not considering alternatives which involve driving a vehicle at all.

$P_n(\alpha \mid \beta)$ means – barring applicability restrictions on n :

$$\neg\alpha \sqcap \beta \sqsubseteq \exists \preceq_n \alpha \sqcap \beta$$

If the permission is for instance *In an emergency situation, drivers of an emergency vehicle are allowed to race a red light if circumstances permit*, the interpretation in terms of subjunctive betterness, involving some interpretation to separate the condition from the thing permitted, is:

When an emergency vehicle stops for a red light while there is an emergency situation and circumstances permit racing the red light, an equal or better alternative is racing the red light.

As pointed out in section 4.6, there is a difference between permissions that offer a free choice between two alternatives, and those that remain silent on the other option. If the phrasing of the text does not suggest such a freedom of choice, one should go for this interpretation. Indifference of the legislator simply involves a permission in both directions, that is adding:

When there is an emergency situation and circumstances permit racing the red light and an emergency vehicle races the red light, an equal or better alternative is that the emergency vehicle stops for the red light.

The legislator may not intend this.

In order to make a number of formal observations about the family of \succeq properties I present it here as if it were a self-contained labeled modal logic. In actuality we are however dealing with simple transitive OWL DL properties. The property \prec and $=$ are pseudo-properties that do not play an actual role in OWL DL.

As pointed out in section 3.4, OWL DL can be understood in terms of modal frames. Let M be a modal frame of the form $M = \langle W, N, \{\preceq_n\}_{n \in N}, V \rangle$ where W is a set of worlds, N a set of preference theories, the \preceq_n are transitive relations, and V is a propositional valuation. Each $n \in N$ is a (consistent) subtheory in the larger preference theory, and can be intuitively understood as a kind of (institutional) agent that imposes its own preferences on worlds $w \in W$. In a description logic context the worlds should be understood as the individuals in a model (cf. the examples in section 3.4.2), and specifically as actions or the alternative plans in a menu.

The preference theory represents the *norm* implied by one normative rule, or if preferred by a set of normative rules provided they are consistent with each other. A set of preference theories describes a normative order. The notion of *preference* and *preference theory* as used in this section is strictly used for compatibility with relevant preference theory literature (generally [138, 241, 267, 100, 102, 101]): what we are

dealing with are subjunctive propositions about *betterness*, made by a legislator, that are intended to play a role as motivating preferences in decisions.

As for instance Sagoff explains (cf. [241]) preferences are generally attributed to *agent* roles (see i.a. section 4.5.2); Human beings cannot be attributed a single comprehensive agent role, and the set of preferences ascribed to them in order to explain their behaviour need not be consistent. The concept of attributing mental attitudes to normative systems as if they were agents has been used before in computer science & Law (cf. for instance [28]). The observation Sagoff makes applies here: certain sets of normative rules may describe a coherent normative order as a preference theory, but the totality of rules of an institution does not have to, and often doesn't.

Hansson in [138] gives a state-of-the-art account of obligation based on preference theories. This is not such an account; The normative rule only imposes constraints on plans when alternatives are considered that are allowed or disallowed by the normative rule. The proposal given here does not allow for evaluation of normative positions in the way that a deontic logic does, but it may be combined with such an account.

Read $x \preceq_n y$ as “preference theory n considers y as least as good as x ”. Read $x \prec_n y$ as “preference theory n considers y better than x ”. This is shorthand for $(x \preceq_n y) \sqcap \neg(y \preceq_n x)$. Read $x =_i y$ as “preference theory n is (explicitly) indifferent between y and x ”. This is shorthand for $(x \preceq_n y) \sqcap (y \preceq_n x)$. In this book the label n is omitted in some cases.

The language L is a set of terminological axioms, implicitly joined by a \sqcap , in a description logic *TBox* (terminological box) of the form $\phi_1 \sqsubseteq \phi_2$ where ϕ_1 and ϕ_2 are concepts conforming to the syntax in 3.4. As usual $\forall \preceq$ is defined in terms of $\exists \preceq$ as follows: $\forall \preceq \phi \equiv \neg \exists \preceq \neg \phi$. The truth definition for this language, and semantic notions like frame, satisfiability, and validity, are standard (see [24]):

$$M, w \models \exists \preceq_n \phi \text{ iff } \exists w' : w \preceq_n w' \text{ and } M, w' \models \phi$$

This says that ϕ is true in at least one alternative of w which n considers at least as good as w . The preference modality $\exists \preceq_n$ constrains a preference order at the level of worlds, for a preference theory n .

Note that the frame axiom for reflexivity, which would allow me to turn \preceq into a reflexive relation, is absent: OWL DL does not allow for reflexive frames. Addition of this axiom would permit use of the *S4* system used by i.a. Van Benthem et al. as a preference logic (cf. [267]), and as a building block for more complex deontic logics like for instance Boutilier's *CT4O* (cf. [56]) or Van der Torre's *2DL* (cf. [268]).

Reflexivity is a common property of mereological, topological, spatial, and temporal relations. Some workarounds have been proposed in i.a. [125] (submissions 26 and 33), and the OWL specifications propose some alternative methods for approximating reflexivity in the definition of parts and wholes. The OWL 2 proposal (cf. section 3.4) does add reflexivity, but is not the standard yet. Acceptation of OWL2 would be reason to reconsider this formalization.

One of the arguments made for the original omission of reflexive frames, besides the apparent computational attractiveness of doing so, is that its primary function is as an abbreviation anyway: instead of stating that for some w there is some w' such that wRw' that has some property we are looking for, we have to state that w has the

property we are looking for or there is some w' such that wRw' that has the property we are looking for. In [234] Rescher made the point that the property of reflexivity is in most use cases actually ontologically suspect: nothing is a “part of itself” and well-formed wholes should consist of at least two parts. In a mereological, topological, spatial, or temporal context this observation is not of great importance. If identity is not regarded as a limit case of parthood this does not appear to invalidate important results.

Rescher’s point clearly also holds for the \preceq relation: it is odd to say that something is as good as itself, or that we are indifferent between something and itself. Since we are mostly concerned with the irreflexive counterpart of \preceq the limitations of OWL DL can be accepted.

It is unfortunate that this limitation of OWL DL necessitates re-evaluation of existing formal results before they can be incorporated into the Semantic Web.

As stated, the obligation $O_n(\alpha \mid \beta)$, or a prohibition $F_n(\neg\alpha \mid \beta)$ translates to:

$$\begin{aligned}\beta \sqcap \alpha &\sqsubseteq \neg\exists \preceq_n \beta \sqcap \neg\alpha \\ \beta \sqcap \neg\alpha &\sqsubseteq \exists \preceq_n \beta \sqcap \alpha\end{aligned}$$

This says that in a world where $(\beta \sqcap \alpha)$ is true, there is no world equal or better accessible where $(\beta \sqcap \neg\alpha)$ is true, and in a world where $(\beta \sqcap \neg\alpha)$ is true, there is some world equal or better accessible where $(\beta \sqcap \alpha)$ is true, according to i . It is possible to include other worlds where $(\beta \sqcap \neg\alpha)$ is true in the ordering as long as one does not claim they are better than a world in which $(\beta \sqcap \alpha)$ is true, and obviously a world where $\neg\beta$ is true can be placed anywhere in the ordering.

Because the peculiar frame-based syntax of OWL will produces a triad C, C', C'' in a taxonomy, as discussed in section 6.1.1, there will often be good reason to simply state the subsumption triples $C' \sqsubseteq C, C'' \sqsubseteq C$, and a disjointness triple $\neg C' \equiv C''$, in addition to the $C' \sqsubseteq \neg\exists \preceq_n C'', C'' \sqsubseteq \exists \preceq_n C'$. In other words, α as differentiating characteristic, between C' and C'' can be left implicit.

The permission $P_n(\alpha \mid \beta)$ translates to:

$$\beta \sqcap \neg\alpha \sqsubseteq \exists \preceq_n \beta \sqcap \alpha$$

This says that in a world where $(\beta \sqcap \neg\alpha)$ is true, there is some world equal or better accessible where $(\beta \sqcap \alpha)$ is true. Intuitively this may suggest that it is advisable to choose $(\beta \sqcap \alpha)$ over $(\beta \sqcap \neg\alpha)$, but that is not intended. It is obviously consistent to add an edge from a world in which $(\beta \sqcap \alpha)$ is true to a world where $(\beta \sqcap \neg\alpha)$ is true, establishing that both alternatives are of equal value as proposed in section 4.6.

Note that the characterization of permission is just to alert us to any *conflicts of disaffirmation* between obligations and permissions. Trying to strengthen the definition of permission within the expressive limits of OWL DL (for instance $(\beta \sqcap \neg\alpha) \sqsubseteq (\exists \preceq_n (\beta \sqcap \alpha) \sqcap \forall \preceq (\alpha \sqcup \neg\beta))$) is not going to solve the problem that we are dealing with an underspecified frame that includes unintended orderings.

The relation \preceq is not complete. Some worlds are genuinely incomparable, and they must be if \preceq is to be a reasonable interpretation of permission. This means that the \preceq

relation does not meet the property of *trichotomy*, requiring that it should be possible to choose between any two alternatives, as is often required of preference relations. Hansson defends the general invalidity of the trichotomy property in [138], but this argument is hardly universally accepted. The relation covered here therefore does not match with the mainstream conceptions of *revealed* preferences as explained in section 2.4.1. As pointed out, limitations of OWL offered a practical reason to reject trichotomy.

Proposition 31 (No trichotomy). Some worlds are incomparable.

Proof. Assume the obligations $O_n(\alpha \mid \top) \sqcap O_n(\beta \mid \top)$. $\mid \phi \mid$ is the set of worlds w such that $M, w \models \phi$. Assume a $x \in \mid \neg\alpha \sqcap \beta \mid$ and a $y \in \mid \alpha \sqcap \neg\beta \mid$. Observe that both $x \preceq_n y$ and $y \preceq_n x$ are inconsistent. \square

Hansson tried to explain deontic logic in terms of *ceteris paribus preferences* [102, 100, 101], i.e. *all else being equal* preferences: these also do not mesh well with revealed preference. Subjunctive betterness as encountered in formal constitutive rules, which are obviously formulated prior to the phenomena they describe as pointed out in section 4.1, has little to do with revealed preferences. The notion that preferences can be revealed is based on the assumption that choices between alternatives are observable, which is highly suspect, as pointed out in section 2.4.1.

To be viable as a deontic knowledge representation system, the system has to meet some requirements. Minimally, the system should capture two properties: that what is obliged should be permitted and the impossible should not be obligatory.

The intuition for the first property is that you cannot be obliged to do something without at the same time being permitted to do that something. The property is usually expressed by way of the following axiom: $O\phi \rightarrow P\phi$. This characterization is not defined in terms of the frame axiom $\forall \preceq \phi \rightarrow \exists \preceq \phi$, but it does enforce it:

Proposition 32. What is obligatory is permitted. The axiom $O_n(\alpha \mid \beta) \rightarrow P_n(\alpha \mid \beta)$ is true.

Proof. Trivial from $((\beta \sqcap \alpha) \sqsubseteq \neg\exists \preceq_n (\beta \sqcap \neg\alpha)) \sqcap ((\beta \sqcap \neg\alpha) \sqsubseteq \exists \preceq_n (\beta \sqcap \alpha))$ and $(\beta \sqcap \neg\alpha) \sqsubseteq \exists \preceq_n (\beta \sqcap \alpha)$. \square

Note that it is perfectly possible to weaken the obligation so that it does not meet this property (to appease the critics mentioned in 6.3.1). The obligation translates to two statements in this system, connected here by a \sqcap . Omitting the preference proposition $((\beta \sqcap \neg\alpha) \sqsubseteq \exists \preceq_n (\beta \sqcap \alpha))$ will do the trick. This is however not a feature I would want to include.

A nice feature of this system is that it does not only make the impossible obligation $O_n(\neg\alpha \mid \alpha)$ inconsistent, but will also do something intuitively meaningful with the unintuitive and meaningless obligation $O_n(\alpha \mid \alpha)$ mentioned by [195].

Proposition 33. The impossible is not obligatory: $\neg O_n(\neg\alpha \mid \alpha)$ is true.

Proof. Observe that \perp follows from $((\neg\alpha \sqcap \alpha) \sqsubseteq \neg\exists \preceq_n (\neg\alpha \sqcap \neg\alpha)) \sqcap ((\neg\alpha \sqcap \neg\alpha) \sqsubseteq \exists \preceq_n (\neg\alpha \sqcap \alpha))$. \square

Proposition 34. The meaningless obligation $\neg O_n(\alpha \mid \alpha)$ gives rise to an incoherent concept $(\alpha \sqcap \neg\alpha)$ which evaluates to the empty set represented by concept `OWL:NOTHING`. Concepts such as this will generally be marked as ill-formed by an OWL aware editor like Protégé.

The propositions sofar do not require a deeper understanding of OWL DL subsumption relations between propositions. Contrary-to-duty situations and normative conflict are also classical tests of the viability of a knowledge representation system for normative rules. In their most straightforward formulation these tests are also passed. In section 6.5.1 we will however see that when we take OWL DL subsumption into account this formalization shows peculiar behaviour, when compared to deontic logics.

The contrary-to-duty situation discussed in section 6.4 present no problem, as is shown with the Chisholm paradox.

Proposition 35 (Chisholm). The sentences $O_n(\alpha \mid \top)$, $O_n(\beta \mid \alpha)$, $O_n(\neg\beta \mid \neg\alpha)$ are only satisfied by the ordering identified by [268].

Proof. If $\mid \phi \mid$ is the set of worlds w such that $M, w \models \phi$, then the desirable ordering is represented as $\mid \neg\alpha \sqcap \beta \mid \prec_n \mid \neg\alpha \sqcap \neg\beta \mid \prec_n \mid \alpha \sqcap \neg\beta \mid \prec_n \mid \alpha \sqcap \beta \mid$. Verify for each set $\mid \phi \mid$ that:

1. that it cannot be empty if some other set $\mid \phi' \mid$ in the ordering is nonempty and $\mid \phi' \mid \prec_n \mid \phi \mid$;
2. that $x \preceq y$ is not possible for any $x \in \mid \phi \mid$ and $y \in \mid \phi' \mid$ and $\mid \phi' \mid \prec_n \mid \phi \mid$.

□

The essence of the Chisholm paradox is of course not that to comply with the obligations you have to choose a world $w \in \mid \alpha \sqcap \beta \mid$. The knowledge representation system should be able to establish that if only a settled context $\mid \neg\alpha \mid$ is given, i.e. all considered alternative worlds $w \in \mid \neg\alpha \mid$, then the optimal choice is a $w \in \mid \neg\alpha \sqcap \neg\beta \mid$.

As mentioned before, the conflicts of disaffirmation and compliance described in detail in section 6.3 do raise a contradiction, as they arguably should. As argued before it is important to be able to calculate all *prima facie* obligations, even if we apply a method to resolve them automatically. The axiom that captures this notion for conflicts of compliance is generally called the *no-dilemma assumption*.

Proposition 36 (No dilemma). There are no conflicting obligations. The obligations $O_n(\alpha \mid \beta)$ and $O_n(\neg\alpha \mid \beta)$ are inconsistent: $\neg(O_n(\alpha \mid \beta) \sqcap O_n(\neg\alpha \mid \beta))$ is true.

Proof. If $M, w \models \beta \sqcap \alpha$ then $M, w \models \exists \preceq_n (\beta \sqcap \neg\alpha) \sqcap \neg \exists \preceq_n (\beta \sqcap \neg\alpha)$ if both statements are true. □

The conflict between a permission and an obligation follows a very similar pattern. As stated earlier the obligation consists of two separate statements. One captures the permission for a choice that it entails as $(x \preceq_n y)$, and the other, negative, statement can be thought of as blocking a permission for a competing choice $(\neg(y \preceq_n x))$.

Proposition 37 (Disaffirmation). The obligation $O_n(\alpha \mid \beta)$ and permission $P_n(\neg\alpha \mid \beta)$ are inconsistent: $\neg(O_n(\alpha \mid \beta) \sqcap P_n(\neg\alpha \mid \beta))$ is true.

Proof. If $M, w \models \beta \sqcap \alpha$ then $M, w \models \exists \preceq_n (\beta \sqcap \neg\alpha) \sqcap \neg \exists \preceq_n (\beta \sqcap \neg\alpha)$ if both statements are true. \square

How to deal with normative conflicts that *do* occur is the subject of the following subsection. For now I will end this presentation with some remarks.

The modal notion of bisimulation makes some sense in this interpretation and can be used to establish that two preference structures are the same from a modal standpoint. This notion is central to the semantics of comparison of preference structures but we will not develop this line of possible research here.

It may be useful in comparing preference structures to include ‘silent’ preferences in analogy to silent transitions in weak bisimulation of state transition systems. As will be pointed out in section 6.7.3, comparison as intended in comparative law involves three systems: the systems M_1 and M_2 to be compared, and the norm(s) of analysis used to decide whether they are sufficiently ‘the same’. It is possible to check whether M_1 is a weak bisimulation of M_2 , but also which one is the “best” simulation of a third set of norms M_3 .

Decidability in LKBS use contexts – if we leave the application of defeasible rules out of consideration – is taken for granted, as it is represented in OWL DL (cf. [158]). OWL DL has been used in practice on very large knowledge bases (viz. [193]), although these exhibit other structural properties than legal knowledge bases built with this method. It is known that certain mereological and temporal constructs do not scale very well, and that is also the case for the betterness relation central to the conceptualization of norms proposed here: it is very taxing on OWL reasoner implementations. The question of decidability of this representation *including* the defeasible rules is ill-defined, since the logic of applying these was intentionally left underspecified.

6.5.1. Betterness and Normative Conflict

In legal theory there is a lively interest in the question what pairs of normative rules – assuming we would a priori expect them to form a coherent normative order – can be in normative conflict. Conflict plays a role on two different levels: there is the true normative conflict on logical grounds, and there are mundane cases where a pair of normative rules presents us with a planning dilemma. Planning dilemmas are not normative conflicts.

Section 6.3 addressed the notion of normative conflict as a relation between two normative rules that are *necessarily* in normative conflict on terminological grounds. In addition, legal practice recognizes conflicts that only occur in certain categories of cases, as section 6.3.1 observed. This latter category can, as a consequence of the observed ontological commitments for the representation of normative rules, not be considered candidate reusable knowledge about sources of law, since the construction of the conflict involves ampliative reasoning of an unspecified nature.

In principle the existence of normative conflict must be affirmed by another source of law to be of direct relevance here. This would be generally case law.

The betterness interpretation of normative rules does provide some useful tools for the detection and resolution of conflicts, which will be discussed in this section.

The first such tool is the epistemic interpretation of what violation is, as an alternative to the straightforward derivation of the institutional qualification.

According to the betterness reading of normative rules, if you choose an alternative even though another alternative was available that is strictly better in n , then you have violated n . This can be interpreted in two ways, depending on whether we use a straightforward model theoretic interpretation of OWL DL or the abox of the OWL DL reasoner as the epistemic model of a theory (as explained in section 3.4.2):

Terminological: the formalization of the normative rule n entails – as a matter of ontology – that there exists a better or equal alternative, and asserting that the alternative a under consideration is (also) better or equal to that alternative (i.e. that they are of equal value) would give rise to inconsistency; or

Epistemic: the abox names some other individual alternative a' , and it is not possible to consistently assert $(a' \preceq_n a)$ (or $\preceq_n(a', a)$ in abox notation).

The terminological option has no added value over the institutional interpretation of normative rules. It is in fact only more complicated than merely establishing the proposition $\exists \text{disallowedBy}.\{n\}$.

The epistemic alternative is important for a variety of reasons. When planning, the comparison between alternatives is generally made between those explicitly considered. The best plan, among those conceivable, is one of which it is possible to assert that it is better or equal than all alternative plans. In addition, in planning we typically deal with settled contexts.

The restriction to a *settled context* relates to the observation in section 6.3.1: we only have to order the alternatives that are known to be feasible. In practical planning there are two reasons to consider certain propositions part of the settled context:

- they are already true in our interpretation of the situation and cannot – to one's knowledge – be undone anyway, or they are inevitably going to become true according to our interpretation of the situation; or
- one already settled on violation of the law, but plans to do so without any other violations.

The violation without a settled context can be understood as a violation in settled context \top .

Proposition 38 (Violation). $u \preceq w$ is a violation of preference theory n in settled context α if and only if $w \prec_n u$ and $M, w \models \alpha$ and $M, u \models \alpha$.

The object of using the preference theory for planning is to order alternatives in such a way that the best alternative remaining does not violate rules, or no additional ones beyond the ones already violated in the settled context.

An important feature of this formalization is that it only takes into account *orderings of alternatives* that are against the rules. It is silent on decision problems where

all available alternatives are against the rules, since these do not involve preferring an alternative that is allowed over one that is disallowed. This distinction however only becomes relevant for planning if an epistemic interpretation is chosen.

In some cases a preference theory based on normative rules may present the planning agent with a planning dilemma.

Proposition 39 (Planning Dilemma). Both $u \preceq w$ and $w \preceq u$ are violations of the relevant preference theory in the chosen settled context.

The planning dilemma is *not* a normative conflict, or significant evidence of it. The planning agent has to choose between violating one of two normative rules. A normative conflict is only deemed to exist if one of the normative rules is not applicable due to the fact that the other also applies.

There is in principle no logical problem with substantially the same behaviour being allowed by some normative rule n_1 and disallowed by n_2 . This inconsistency could only arise if one makes the following assumption about a *set* of normative rules N even though N contains conflicting rules such as the pair $\{n_1, n_2\}$: if behaviour b is disallowed according to some $n \in N$ then behaviour b is also disallowed by N , and if behaviour b is allowed according to some $n \in N$ then behaviour b is also allowed by N .

Some *collections* of normative rules may be expected to result in a coherent account of normative order in terms of a preference theory partially ordering alternatives. If this account is inconsistent, it exposes normative conflicts. If the explanation of a normative order as a single set of equally important propositions fails for the analyzed collection, then an alternative conceptualization of the collection must be found. The construction of alternative conceptualizations of coherent normative order is however beyond the scope of this section: it is addressed in the next one. What is of interest here is how one knows that there is a normative conflict in a set of rules.

The strategy to be used is composing a single compound normative rule out of the set of rules, and investigating the resulting preference theory.

Proposition 40 (Compound Normative Rule). Let n be any normative rule, and n_c be a compound normative rule. For any $n \in n_c$, if behaviour b is disallowed by n then behaviour b is also disallowed by n_c , and if behaviour b is allowed by n then behaviour b is also allowed by n_c . The following must be an axiom of any coherent compound normative rule n_c : $\{n_c\} \equiv \forall \text{allows}.\forall \text{disallowedBy}.\neg\{n_c\}$.

The membership criterium for the set of rules one is interested in is of course application-specific. Let us for instance assume that the normative rules represented by work w , which is a source of law, is an attempt to formalize a coherent normative order. Construct a compound normative rule n_w that combines the normative rules in w :

$$\exists \text{disallowedBy}.\exists \text{representedBy}.\exists \text{metalex:embodies}.\{w\} \sqsubseteq \exists \text{disallowedBy}.\{n_w\} \quad (6.1)$$

$$\exists \text{allowedBy}.\exists \text{representedBy}.\exists \text{metalex:embodies}.\{w\} \sqsubseteq \exists \text{allowedBy}.\{n_w\} \quad (6.2)$$

If one would want to devise a strategy to construct compound normative rules automatically, one would use indicative rules with a ($\text{free}.\exists\text{disallowedBy}.\{n\}$) condition.

If making these assertions, in addition to the coherency axiom, does not introduce inconsistency, then there is no normative conflict within the set.

Proposition 41 (Terminological normative conflict). If a compound normative rule contains terminological normative conflict, it will not be coherent.

Finding *pairs* of rules that pose normative conflict is a diagnostic problem, that can be solved with a systematic approach to composing hypothetical compound normative rules in an OWL DL description classifier. This is done offline, i.e. not in the production LKBS, since the results of this computation are a static knowledge source to the LKBS. The result of such a diagnosis can be stored in the form of choice rules, discussed in section 6.6.3. Choice rules constructed using such a method do not however reflect a source of law, and are of no direct interest to this book.

The terminological normative conflict criterium as considered here does not however cover all types of normative conflict discussed in section 6.3.

Logical inconsistencies are obviously always an artifact of the chosen formalization. Certain classes of deontic logics can be for instance be classified by whether they have axioms for *strengthening the antecedent* ($O(\alpha_1 \mid \beta_1)$ to for instance $O(\alpha_1 \mid \beta_1 \sqcap \beta_2)$) and *weakening the consequent* ($O(\alpha_1 \mid \beta_1)$ to for instance $O(\alpha_1 \sqcup \alpha_2 \mid \beta_1)$) (cf. for instance [268] for an overview of deontic logics categorized by these features). Depending on what strategy one chooses, one is more or less likely to find logical contradictions between normative propositions.

These features always attract criticism. Certainly in law, where the presence of normative conflict is a matter to be explained, the absence of normative conflict is often considered a feature of legal deontic logics. The normative conflict however only arises from the deontic interpretation of normative rules; Normative conflict can in that sense be considered an essential feature of deontic logic proper. Because the formalization given here is only a basic *ordering* one, which treats the antecedent of the obligation very differently from deontic logics, a transparent comparison with deontic logics that can be classified as having the features of *strengthening the antecedent* or *weakening the consequent* is not trivial. The behaviour of the antecedent is best illustrated by an extended example.

Example 13. Consider a simplistic example: Assume that $\text{MOWLAWN} \sqsubseteq \text{ONLAWN}$ (hence $\neg\text{ONLAWN} \sqsubseteq \neg\text{MOWLAWN}$) and $\text{TUESDAY} \sqsubseteq \text{WORKDAY}$ are true.

Consider the following disaffirmation conflict: It is disallowed to walk on the lawn ($F_n(\text{ONLAWN})$ or $O_n(\neg\text{ONLAWN})$), but it is allowed to mow it ($P_m(\text{MOWLAWN})$). Alternatively consider the compliance conflict version: It is disallowed to walk on the lawn ($F_n(\text{ONLAWN})$ or $O_n(\neg\text{ONLAWN})$), but it is *obliged* to mow it ($O_m(\text{MOWLAWN})$).

Both pairs of two normative rules become logically inconsistent in the chosen representation if we would consider \preceq_m and \preceq_n as being part of a more general betterness relation \preceq_{nc} :

$$\neg \text{ONLAWN} \sqsubseteq \neg \exists \preceq_n \text{ONLAWN} \quad (6.3)$$

$$\text{ONLAWN} \sqsubseteq \exists \preceq_n \neg \text{ONLAWN} \quad (6.4)$$

$$\text{MOWLAWN} \sqsubseteq \neg \exists \preceq_m \neg \text{MOWLAWN} \quad (6.5)$$

$$\neg \text{MOWLAWN} \sqsubseteq \exists \preceq_m \text{MOWLAWN} \quad (6.6)$$

Axiom pairs 6.3 and 6.6, *and* 6.4 and 6.5 are inconsistent with a theory of brute reality that recognizes that $\text{MOWLAWN} \sqsubseteq \text{ONLAWN}$ and $\neg \text{ONLAWN} \sqsubseteq \neg \text{MOWLAWN}$. Subsumption between consequents appears to do the expected job here, for both cases.

The addition of an antecedent (for instance $O_n(\neg \text{ONLAWN} \mid \text{WORKDAY})$) makes no difference as long as both normative rules have the *exact same* antecedent.

Pairs of the type $O_n(\neg \text{ONLAWN} \mid \text{WORKDAY})$ and $P_m(\text{ONLAWN} \mid \text{TUESDAY})$, where the antecedents subsume eachother, but the consequent is exactly the same also make no difference. Subsumption between antecedents also does its job, and as expected results in a reversal

Pairs however cannot vary in both the antecedent *and* consequent. Consider the axioms from the example above, but with the condition WORKDAY or, more specific, TUESDAY added, resulting in $O_n(\neg \text{ONLAWN} \mid \text{WORKDAY})$ and $O_m(\text{MOWLAWN} \mid \text{TUESDAY})$:

$$\text{WORKDAY} \sqcap \neg \text{ONLAWN} \sqsubseteq \neg \exists \preceq_n \text{WORKDAY} \sqcap \text{ONLAWN} \quad (6.7)$$

$$\text{WORKDAY} \sqcap \text{ONLAWN} \sqsubseteq \exists \preceq_n \text{WORKDAY} \sqcap \neg \text{ONLAWN} \quad (6.8)$$

$$\text{TUESDAY} \sqcap \text{MOWLAWN} \sqsubseteq \neg \exists \preceq_m \text{TUESDAY} \sqcap \neg \text{MOWLAWN} \quad (6.9)$$

$$\text{TUESDAY} \sqcap \neg \text{MOWLAWN} \sqsubseteq \exists \preceq_m \text{TUESDAY} \sqcap \text{MOWLAWN} \quad (6.10)$$

Obviously this pair of normative rules will present us with dilemmas with respect to a certain class of behaviours in which TUESDAY is the case, but the two norms are not *necessarily* in conflict: there is no *logical* inconsistency in the chosen formalization. Neither does the reversal of WORKDAY and TUESDAY result in inconsistency.

Proposition 42. A terminological normative conflict between a pair of normative rules only exists if either the antecedents or the consequents of the deontic propositions that were translated to normative rules are equivalent.

The conclusion must be that the given formalization finds only a subset of the types of normative conflict discussed in section 6.3. It differs substantially from the informal presentation based on sets representing the extension of concepts used in that section, resulting in an important deficiency. It however does so on the safe side: the normative conflicts found are definitely to be considered normative conflicts.

The most obvious way to address this deficiency is to expand the set of normative rules in the compound normative rule, by inventing new ones. Let a pair of normative rules n, m of interest for instance express $O_n(\neg \text{ONLAWN} \mid \text{WORKDAY})$ and $O_m(\text{MOWLAWN} \mid \text{TUESDAY})$; Given $\text{TUESDAY} \sqsubseteq \text{WORKDAY}$ it makes sense to conjecture that a rule $O_n(\neg \text{ONLAWN} \mid \text{TUESDAY})$ exists, based on n . Among the

many arbitrary – and defeasible – conjectures one could make of this kind, this one is interesting because of the existence of m .

Given this deficiency – which does not affect practical planning and dilemma in planning – there is still a case to be made for adding a deontic logic on top of this proposed interpretation for the purpose of discovering more normative conflicts.

In summary, this section constructs two general characterizations of the problem of conflict or dilemma. A planning dilemma exists if two alternative actions are ordered in opposite directions by two normative rules. If only the alternative actions in the abox are considered, one obtains an epistemic interpretation of violation and dilemma. There is also a more specific criterium for true normative conflict: two normative rules are intended to be part of one single coherent normative order, but cannot be conceived of as a single compound normative rule. The normative conflict necessarily exists between both normative rules. Somewhere in between these two criteria is a class of normative conflicts that are only normative conflicts for a specific category of cases, and, as far as we should be concerned, and *only when a source of law states it is*.

Sources of law express resolutions of normative conflicts in the form of choice rules. These are introduced in the next section.

6.6. Choice between Legal Rules

We all know that beliefs can sometimes be wrong, so intelligent beings need to be able to revise beliefs when they acquire new information that contradicts their old beliefs. Reasoning systems modeling this phenomenon are called *belief revision* systems.

One common way of determining which beliefs should be surrendered is to use a so-called *epistemic entrenchment* ordering (viz. [208]) discussed in section 2.4. This ordering expresses the idea that some of our beliefs are more fundamental than others; It is a “betterness” ordering on beliefs.

This preference ordering is distinct from the betterness ordering that describes a normative order. The preference ordering on beliefs guides us in surrendering beliefs and adopting new beliefs as our understanding of a situation improves. The preference ordering alluded to by the normative order orders situations and actions and guides us in determining whether an action is desirable, and how a situation is to be avoided or reached.

Legal reasoning, like most domains of reasoning, involves both types of ranking the better over the worse. Since both types of reasoning are often successfully assumed to be independent, i.e. no deontic preference ordering ever follows from an epistemic preference ordering and vice versa, we often limit our attention to each type of reasoning individually (as pointed out in section 2.4.1).

The mapping between brute and institutional realities occasionally turns out to be less than perfect. Our beliefs about brute reality are generally speaking epistemically deeper entrenched than our beliefs about institutional reality: we generally assume, in the context of LKBS, that institutional facts only exist as long as it is consistent to believe they exist. In the case of law this means the following: If application of the

whole set of rules leads to absurd results, we must obviously choose not to apply one of the rules.

6.6.1. Lex Posterior, Specialis, and Superior

Makinson and Gardenförs postulated in [117] that there is a tight connection between belief revision and nonmonotonic logics. Belief revision leads to *temporal nonmonotonicity*, i.e. the set of beliefs does not grow monotonically with time. Default reasoning leads to *logical nonmonotonicity*, i.e. the set of consequences does not grow monotonically with the set of premises.

In law we find a parallel process: we assume that over time legal institutions refine the rules as they get better at capturing the intended normative order. In legal theory we find the principles of *lex posterior derogat legi priori* (*lex posterior*) and *lex specialis derogat legi generali* (*lex specialis*) (cf. generally [265, 189]).

The *lex posterior* principle entails that, in case of irreconcilable conflict, the later legal rule will take precedence over the earlier legal rule, and the *lex specialis* principle entails that, in case of irreconcilable conflict, the more specific legal rule will take precedence over the more general legal rule. There is also the third principle of *lex superior*, or *lex superior derogat legi inferiori*, which states that the higher legal rule will take precedence over the lower legal rule (cf. generally [265]).

The *lex specialis* and *lex posterior* principles *describe* certain phenomena of reasoning in general. They do not sanction the preference of the newer over the older, or the specific over the general, but merely observe that it usually *is* so.

The *lex specialis* and *lex posterior* principles do not have to be explained in many words to laymen; They will be naturally applied even by children in contexts outside law. The *lex specialis* principle, as a kind of logical nonmonotonicity, is based on the principle of parsimony in communication (cf. section 2.5). We expect the reader or listener to infer – to some degree – that our more specific statements are *exceptions* to the more general ones if they appear to be in conflict. The *lex posterior* principle, as a kind of temporal nonmonotonicity, is based on the assumption of *improvement*: if later and earlier statements from the same source appear to be in contradiction, the reader or listener will generally speaking assume that the later statement reflects a better understanding of the issue by the author than the earlier statement, and that the author of the statement is aware of the fact that he is revising an earlier statement. We take for instance for granted that the latest publication discussing a theory reflects the best understanding of that theory.

As I explained in [48], the situation is more complicated in law when reasoning over cases that happened in the past. This situation is discussed in the context of version management in section 5.3. In normal circumstances we would never consider applying the scientific theories of the 10th century to the reconstruction of an event in the 10th century, but in law we are often asked to apply the rules of the past to a case of the past.

Logical and temporal nonmonotonicity have arguably become explicit in law, as the legal principles of *lex specialis* and *lex posterior*, because law can manipulate these principles by for instance instructing the reader to not apply them. It is exactly because

legal reasoning occasionally deviates from the intuitions embodied by *lex specialis* and *lex posterior* that in the context of law people will start doubting their common sense judgment.

Conceptually, application of the principles is easy to understand. Any provision obviously has a history – when did it enter into the norm system and when where changes last made to the text – and it has propositional content. A *lex specialis* ordering between two provisions can be discovered by comparing the provisions content-wise. A *lex posterior* ordering between two provisions can be discovered by comparing the history of the provisions.

The *lex superior* principle is also ‘intuitive’ enough, but it is a design principle for complex, layered legal systems. An act of parliament for instance typically takes priority over a royal decree, which takes priority over a municipal decree, etc. This hierarchy is not something which is discovered by application of the principle to a pair of provisions. It has been purposefully designed into the legal system by a legislator who foresaw the possibility of having to choose between the rules. Provisions overruling the normal activity of *lex specialis* and *lex posterior* usually define a preference ordering on sources of law that can be used for choosing between norms. The *lex superior* ordering is a consequence of the provisions that manipulate the application of *lex specialis* and *lex posterior*.

It is only natural to wonder whether legislators cannot invent additional principles beyond the three mentioned. These would however all be classifiable as instances of *lex superior* choice rules. An odd example (*viz.* [260]) is for instance to leave the choice which rule to apply to the defendant. The implementation of the *lex superior* principle as a rule would in this case be dependent on the decision making setting, but still be classifiable as an instance of *lex superior*.

The logic of the legal system dictates that *lex superior* should take precedence over *lex specialis* and *lex posterior*. The relative priority of the *lex posterior* and *lex specialis* principles among themselves is however not necessarily settled. The reason for this is that they are implicitly assumed to reinforce each other. Assuming that the legislator refines his expressed preferences over time, is aware of his own acts in the past, and intends it’s new provisions to be compatible with the existing corpus, it is only reasonable to expect that new preference statements refine the existing system. The *lex posterior* ordering and the *lex specialis* ordering are in other words expected to point the same way in most cases, while the *lex superior* ordering is expected to point in the opposite way.

In the case where an older rule is more specific than a newer rule, the older rule seems to amend the newer one – which is absurd – and there are good arguments for precedence of both principles: the legislator could have repealed the older rule but didn’t, and the older rule is not likely to have been intended as an amendment to the first one. To the author the first argument is more convincing, and therefore *lex posterior* should defer to *lex specialis*⁴. The law however claims priority over what we consider common sense: the status of this type of conflict has to be verified for each case where it is encountered.

⁴But [266] for instance takes the opposite point of view.

Absurd is the case of a symmetric disaffirmation conflict where even the *lex posterior* principle doesn't offer a solution because the involved provisions were adopted at the same time. This case cannot be solved by interpreting the intent of the legislator, since a rational legislator cannot possibly have intended to contradict himself. In the Netherlands this case would for instance be handled by a detailed provision⁵ that proposes that the order in which the monarch signs⁶ and the sequential order of publication is decisive.

Any criterion will do: For example, one U.S. court held that when two amendments adopted at the same time are irreconcilable, the one receiving the greater number of votes in the Senate takes priority over the other (*viz.* [260]). The absurdity of this situation becomes clear if one realizes that at least some senators voted to support two amendments with the exact opposite propositional content. This exotic principle has not been accepted in jurisprudence anywhere (including the US).

Typical for law is however that there is assumed to be some rule that decides which rule to choose in the case of conflict between a pair of rules. There is no freedom of choice in the matter. Once such a principle has been decided on for a pair of rules, it is subject to *stare decisis* and forms a precedent, an *epistemic obligation* – just like any other rule – to decide in conformance with the new rule.

6.6.2. *Ne bis in idem*

We do not however only have to choose between rules if they are in conflict. There is another legal principle that can also be conceived of as having to choose between rules that does not involve a conflict.

In this case we have to choose between the application of two rules even though no inconsistency, nor a circularity between alternatives, has arisen.

Affirmation of obligations (*cf.* 6.3) does not generally require a choice between the involved norms, but there are real world situations where this choice is demanded by law. In criminal law it is a generally accepted rule that you cannot be prosecuted twice for substantially the same facts. Since there may be several crime descriptions involving for instance unlawful killing, coercion, or defamation, one will occasionally be required to choose between norms with an obligatory character in this area of law.

Ne bis in idem means that no legal action can be instituted twice for the same cause of action. Closely related is *double jeopardy* in common law jurisdictions, which is limited to criminal law. The reason why we have to choose between rules in this case is that the institutional facts created by these two rules cannot be jointly applied to the same thing because doing so would be unjust. This rule is however often limited to criminal law: it is perfectly valid to prosecute for reckless driving while giving an administrative fine for speeding for substantially the same behaviour, provided that speeding is part of another domain of law exempt from the *ne bis in idem* rule vis-à-vis criminal law. The *ne bis in idem* principle only functions in areas of law where it is

⁵ *Aanwijzingen voor de Regelgeving* 1990, art. 173a.

⁶ Contrary to what one might think this order is no longer unique: certain decisions with the status of royal decree, naturalization decrees, are signed once per batch. Apparently no problems are expected here.

specifically declared as a positive principle: it is for instance found in the European Human Rights Convention.

The *ne bis in idem* principle usually applies to an affirmation relationship between two normative rules. For instance insult is an abusive attack on one's honour or dignity, and defamation is an insult involving a false accusation or malicious misrepresentation. In Dutch criminal law one can be prosecuted for either one or the other, but not both. The Netherlands is one of the few jurisdictions that (still) consider these behaviours crimes in addition to considering them torts; A criminal prosecution does not however prevent a (civil) tort action.

Nearly universal is the distinction between murder and manslaughter, or even more grades of increasing seriousness of responsibility for someone's death. There appears to be a kind of informal choice rule that says that one usually prosecutes for the offense with the greater punishment, and this rule often happens to coincide with *lex specialis*: the crime with the more specific description is the one punished more. This suggests an implicit order $\text{DEFAMATION} \succ \text{INSULT}$, $\text{MURDER} \succ \text{MANSLAUGHTER}$.

This is however a prosecution policy which arises naturally from the personal preferences of public prosecutors and victims and not a general principle. If someone is for instance the victim of a defamation, but establishing that this is the case would bring additional harm⁷, the victim will generally prefer a charge with insult instead of a defamation and has the freedom to choose in his own interest.

In this case the operative principle is that the decision maker generally has a freedom of choice between the rules, unless the law states otherwise. It therefore also means that a constitutive *legal act* is performed. If the freedom of choice is restricted by a specific rule, one is (epistemically) obliged to decide in conformance with the rules.

6.6.3. Representation of Choice Rules

The issue of choice between legal rules only comes up for indicative rules. Indicative default rules are applied only, producing a result, if it is consistent to do so, and in some cases only if one intends to bring about that result. Terminological and requirement rules necessarily apply.

If application of an indicative rule makes the knowledge base inconsistent, we retract a some default consequent of an indicative rule and decide that the indicative rule that sanctions it not applicable. Normally speaking, an important epistemic constraint regulates this revision activity in law: the revision must be consistent with choice rules.

Choice rules impose a kind of partial order over defaults by making the applicability of legal rules depend on the application of other legal rules. Choice rules are applicability rules, and just like all other legal rules discussed in this book, they are based in a source of law.

⁷Take for instance the scenario of a married man who visits prostitutes, and is publicly accused of visiting underage prostitutes. Regardless of who has the burden of proof, establishing the truth of the matter will bring additional harm.

In the previous sections a number of different perspectives on choice between rules, and a number of different *reasons* to do so were discussed. There is difference between choice between two indicative rules:

1. because of an inconsistency in institutional reality if both are applied to the same fact;
2. because of incoherence in the intended normative order ascribed to the legislator as evidenced by normative conflict or the dilemma, if both indicative rules are normative; and simply
3. because they cannot be applied to the same fact according to the rules or a decision making policy, even though doing so would lead to no inconsistency or incoherent normative order.

That application of both rules would lead to inconsistency is not of direct relevance to the choice: if the choice rule says that two rules cannot be both applicable, then they necessarily are not. No inconsistency or conflict is required as a condition.

The application of normative rules was implemented in section 6.2 by the standard default rules that are also used for other indicative rules. A pair of rules that is in normative conflict will however in the normal interpretation of normative rules as constitutive rules that derive the institutional fact *allowed* or *disallowed* with the rule itself as axis of reference never be inconsistent with each other: there is in principle no logical problem with substantially the same behaviour being allowed by r_1 and disallowed by r_2 . This inconsistency would only arise if one makes an assumption about a set of rules R even though R contains conflicting rules: if behaviour b is disallowed according to $r \in R$ then behaviour b is also disallowed by R . Set R can be conceived of as a compound normative rule. The underlying assumption, disproven by the verifiable existence of normative conflict, is that R , conceived of as a simple set of rules, formalizes a coherent normative order. If this assumption does not hold, one can retreat to the position that set R , which cannot be represented by a compound normative rule, *and* a set of choice rules formalizes a coherent normative order.

As pointed out in section 3.3, prioritized default reasoning with indicative rules and OWL DL can be conceived of as the use of an *incision function* for belief base revision. The purpose of the incision function is to specify an epistemic entrenchment theory on beliefs, to decide which belief should be retracted in case of inconsistency. Choice rules do not define such an incision function, but constrain it by imposing a noncircular partial order $\{N, \prec\}$ on rules. How deeply an institutional belief is entrenched is determined by the rule that generated it. Even a comprehensive catalogue of such ordering constraints would still leave freedom in the construction of coherent models of institutional reality for the same case that can i.a. be used for tactical purposes in argumentation. The partial order does not replace a sound incision function.

The notion of a noncircular partial order on legal rules representing their priority, and propositions that describe this order is common in computer science & law. See for instance [133, 179, 266, 286, 230] for examples of use of this explanatory device. There are however considerable differences in approach.

[266, 286] for instance only apply it to normative rules, and develop, for normative rules of limited expressiveness only, a method to compute the application of *lex*

specialis to normative conflict (which is criticized in section 6.3.1). The method of [286] subscribes to the notion that a pair of rules must be necessarily in conflict: the partial order on rules can therefore be completely compiled out in advance using a terminological concept of conflict, and the principles of *lex specialis*, *lex posterior*, and *lex superior*. In [266] the computation of normative conflict depends on the case (cf. section 6.3.1), and conflicts are resolved as they are found. A property that both approaches share is that a determined order of priority between two rules is *never* conditional on anything.

The accounts found in [133, 179, 230] potentially apply to any pair of legal rules, and they moreover allow for the possibility that priority between two rules depends on the setting in which the priority relationship is determined because a choice has to be made. This allows for rules that determine the priority between two legal rules for a specific category of cases, which was suggested as a conceptualization of *intersection conflict*.

Neither of these approaches makes a conceptual distinction between “normal” defeasibility of legal rules and the resolution of perceived normative conflict, and they moreover all depend on the notion that choice only follows from conflict or inconsistency.

Hage and Verheij (i.a. [133]) also take a strong position with respect to the status of (court) decisions in general, including the decision to give one rule priority to the other. Such decisions in their opinion cannot have the status of a legal rule. Principles such as *lex specialis* can be said to be part of positive law, but the application of the principle in a decision is not. Here we do not adopt this point of view, which is rather philosophical: for LKBS development a court decision that explicitly orders a pair of rules will always be considered safer ground than the opinion that *lex specialis* applies to them and imposes a certain ordering. This issue is addressed in more detail in section 6.7.2.

If one takes a cautious approach to representation of sources of law one does not try to precompile a partial order based on theoretical arguments on how such priorities arise, but simply represents the ordering as it is made explicit over time by actual decisions and case law, at least by the ones which of one believes that they have a *stare decisis* effect. This is not automatically the case when freedom of choice is involved, like in the *ne bis in idem* case, or when a decision maker chooses between rules to resolve a compliance dilemma (*intersection conflict*). For a reusable representation of sources of law there is of course a good argument to be made for this. Being overly cautious however leads to ignoring the underlying normative order, and makes legal institutions fundamentally incomparable on the functional level, as explained in section 6.7.

The representation solution chosen here is in the tradition of Kowalski in [179], who came to a similar solution for applicability propositions as this book, and treat choice rules as applicability rules.

The incompatibility between pairs of legal rules as applied to the same behaviour can be enforced by applicability rules. If we for instance say that rule r_1 is defeated by r_2 when applied to the same fact because r_2 is more specific, this fact can be represented as *rule r_1 is only applicable if r_2 is not applicable, according to r_s* :

$$\{r_1\} \sqsubseteq \forall \text{appliesTo}.(\forall \neg \text{applicable}.\{r_2\} \sqcap \exists \text{applicable}.\{r_s\})$$

Rule r_s is not the general principle that justifies the inference, but the specific institutional fact – ideally backed up by a source of law that represents it – that r_1 is only applicable if r_2 is not applicable.

The attractiveness of considering the choice principles as generators of applicability rules is that a generous amount of expressiveness is available. It is for instance easy to state that rule r_3 is defeated by r_4 when applied to civil servants because r_4 is more specific when applied to civil servants:

$$\{r_3\} \sqsubseteq \forall \text{appliesTo}.(\neg \exists \text{actor}.\text{CIVILSERVANT} \sqcup (\forall \neg \text{applicable}.\{r_4\} \sqcap \exists \text{applicable}.\{r_{cs}\}))$$

This immediately gives us a solution for cases of *intersection conflict* in which we believe that a specific decision to choose between the rules *will* have effect on future cases, but only in relation to a specific category of cases, as considered in section 6.3.1. Compare for instance dynamic priority in [230]); Dynamic priority cannot be solved by a straightforward partial ordering $\{N, <\}$ on legal rules only.

Constraints of these type enforce the incompatibility of pairs of rules, but they do not enforce an order for an *incision function* (cf. section 3.3). They perform the ontological function of constraining which models are models of institutional reality. Since application of an indicative rule is an ampliative operation it is consistent to say that r_1 is applicable as long as no one has applied r_2 . These rules indicate that two rules cannot be both applied to the same thing, because of general defeasibility, normative conflict, or otherwise.

What we want to encode are statements of the form *applying r_2 would be better than applying r_1* , or *applying r_2 when the case is about a civil servant would be better than applying r_1* . So far we have however always translated what could be considered an *epistemic obligation* – an instruction on “how to think” – to straightforward axioms, constraints, or default rules. Only in the case of obligations relating to behaviour we consider it necessary to make a distinction between compliance and violation.

If we would want to represent this meta-level domain of applying rules, and consider application of an indicative rule an action, we can simply represent them in the same way as we would represent other normative rules, in a new meta-level domain that consists of a case, facts of the case, rule applications, in a specific order, and resulting facts. No LKBS would however ever take this approach, and so far we have settled on an approach which takes `applicable(CONSTITUTIVETHING,RULE)` as the representative primitive and not the reified `APPLICATION` in relation to a `RULE` and `CONSTITUTIVEFACT`.

In the *LKIF* ontology (cf. [35]) I have provided the ingredients for this approach. Assuming that we commit to doxastic-style beliefs, we can consider alternative beliefs one could hold with respect to the same thing. This is basically the same approach as the ordering of behavioural alternatives. If a proposition **proposes** the existence of something, then the `LEGALRULEAPPLICATIONPROPOSITION` proposes something of the form `∃applicable.LEGALRULE`. The proposition that proposes this is a different

object from the one about which this belief is held. The choice rule then becomes a standard normative rule applicable to $\exists \text{proposes}.\exists \text{applicable}.\{r_1, r_2\}$ ⁸:

$$\begin{aligned}\exists \text{proposes}.\exists \text{applicable}.\{r_1\} &\equiv \neg \exists \text{proposes}.\exists \text{applicable}.\{r_2\} \\ \exists \text{proposes}.\exists \text{applicable}.\{r_1\} &\sqsubseteq \exists \preceq_{r_s} . \exists \text{proposes}.\exists \text{applicable}.\{r_2\} \\ \exists \text{proposes}.\exists \text{applicable}.\{r_2\} &\sqsubseteq \forall \preceq_{r_s} . \neg \exists \text{proposes}.\exists \text{applicable}.\{r_1\}\end{aligned}$$

Any observation made about normative rules also applies here, including the problem of normative conflict. Note that the CIVILSERVANT could also be added effortlessly, as well as conditions on the believer in or the medium of the proposition.

The presence of \preceq between two beliefs is sufficient to make clear that they are alternative explanations about the same underlying thing.

Of course, if one does not take a cautious approach to representation of sources of law one does want to try to precompile a partial order based on theoretical arguments on how such priorities arise, as – amongst others some of whom were mentioned earlier – I have done in [286]. For one thing one obviously wants to express known lex superior orderings that apply to whole classes of rules R_1 and R_2 . Because we do not know that any pair of these cannot be applied together – it is even extremely unlikely – we cannot apply them as explained above. We *can* however ensure that no ordering is made that goes against lex superior, and – as explained earlier – lex superior must necessarily take priority: the lex superior ordering is always *allowed*:

$$\exists \text{proposes}.\exists \text{applicable}.R_1 \sqsubseteq \exists \preceq_{r_s} . \exists \text{proposes}.\exists \text{applicable}.R_2$$

Since these lex superior propositions on whole classes of legal rules do explicitly occur in sources of law, the above could very well be a real legal rule.

The same is not true for lex specialis and lex posterior. These are, as stated before, descriptions of the strategies used in defeasible reasoning in general.

Lex specialis causes specific problems: it is not always clear which one of two rules is the more specific one, certainly if the two rules do not logically contradict each other. Specificity suggests some relation with logical subsumption, but two things that have a subsumption relation cannot be in conflict.

Although attempts to formalize the meaning of lex specialis, such as in [286, 265], are certainly interesting from a legal theoretic point of view, automated application of the lex specialis principle has only little relevance to the representation of sources of law. If lex specialis is only the recognition in the legal field of a pre-existing set of generic purpose epistemic habits that the readers of texts bring with them, as argued before, then the underlying mechanism is a more general one involved in text understanding, i.e. decoding information contained in the text.

This mechanism appeals to common sense knowledge – including knowledge with an ampliative function – that the readers bring to the text, and the conclusion of the

⁸That we have asserted $r_1 \neq r_2$ becomes essential here.

application of the principle is relativized to a context if the argumentation involves ampliative inference.

For a representation of a source of law that is used in LKBS in a variety of settings, only the representation of the conclusion of the application of the principle is relevant, and then only if it represents the policy of the LKBS user. In section 6.7.2 a case will be made for the thesis that setting such policies is a major quasi-legislative function of the courts.

6.7. Elaborating on the Intended Normative Order

As pointed out in section 4.7 the objective of this book is to represent legal rules while making minimal epistemic commitments to the way of using these legal rules. At the same time we have to account for two primary uses of legal rules, being planning and plan recognition as pointed out in section 4.1.

Planning involves generating and comparing alternative plans, the execution of which involves behaviour, while situation and action recognition involve generating and comparing consistent alternative explanations of observed behaviour. Both are subject to legal rules: normative rules directly address planning, while all legal rules directly address explaining behaviour in legal terms. Sometimes they have the role of constraints, ruling out certain alternative plans or explanations, sometimes they create viable alternatives, or at least make alternatives that would otherwise be extremely unlikely or undesirable more salient because the behaviour of other agents has become more predictable. Generally speaking, the objective of planning has little to do with the legal rules: law plays only a minor role in people's lives.

Sometimes a rule can be applied if one intends to bring about a certain legal result by proposing that one is executing a plan to that effect (one undertakes a *constitutive act* as explained in section 4.2; One tries to construct a plan that meets $\exists \text{executes}.\exists \text{causes}.\exists \text{initiates}.\text{INTENDEDLEGALPOSITION}$). This effect, intended by the legislator, remains unexplained if we do not consider the larger normative order intended by the legislator.

Chapter 5 proposes a representation of all legal rules, but only the function of normative rules has been elaborated on. By qualifying certain propositions as allowed or disallowed, the legislator attempts to make behaviours occur more, respectively less, often than they did before the normative rule was created. The legislator articulates a preference, in the form of a rule, with the intent that it will motivate the actions of its addressees. The legislator in the Netherlands for instance has publicly announced that he prefers us to drive at the right hand side of the road: underlying this rule are presumably preferences for 1) drivers driving on the same side to prevent accidents and generally increase the capacity of the roads, as well as the speed at which drivers can safely drive it, 2) drivers not having to come to a consensus whether we should keep to the left or right, and 3) not having to switch road side when you drive into a neighbouring country.

Normative rules have an important function for bystanders, regardless of whether the allowed behaviours are also preferred by the bystander: because the legislator is generally successful in making disallowed behaviours occur less often, the bystander

expects that the allowed alternative is much more likely than the disallowed alternative, therefore reasons that it should therefore be (epistemically) preferred as a prediction, and decides based on that assumption. This applies as much to the driver who expects the other driver coming from the left to yield as it applies to the driver who has to yield and who expects that failure to do so may firstly result in an accident, and secondly in additional negative consequences for him made salient only because of the existence of rules that facilitate organizing punishment of his violation.

Non-normative rules are only relevant to situation recognition, with or without a following planning stage that uses the normative rules. Strictly speaking they only mean something in relation to their capacity to trigger normative rules.

The notion that the normative rules are supposed to form a coherent normative order involves an abstraction (to a form of [183]’s up-down metaphor): the \prec_n relation is a relation between domain $|\exists \text{disallowedBy}.\{n\}|$ and range $|\exists \text{allowedBy}.\{n\}|$. Taking the normative order intended by the normative rules, including the associated choice rules, at face value, decision problems involving the normative order can be classified into the following four categories⁹:

Assessment: Does an action a violate normative order \mathcal{N} ?

Envisioning: What categories of cases does normative order \mathcal{N} distinguish, and how does it order them?

Ordering the alternatives: Given a settled context description \mathcal{C} , and a menu \mathcal{A} , and an $a_1 \in \mathcal{A}$, $a_2 \in \mathcal{A}$, order a_1 and a_2 in accordance with normative order \mathcal{N} .

Choosing the best alternative: Given a context description \mathcal{C} , and a menu \mathcal{A} , choose all $a_1 \in \mathcal{A}$ for which there exists no $a_2 \in \mathcal{A}$ such that $(a_1 \prec a_2) \in \mathcal{N}$.

The two differentiating features are whether one is using the normative system with or without a settled context, the situation features present in each alternative, and whether one is ordering or minimizing alternatives.

Assessment can be taken to be a variety of choosing the best alternative (an action violates the normative order if it is *not* in the set of best alternatives). A major difference however exists between the concept *disallowed because the normative order implies that there must be a strictly better alternative* and *disallowed because you chose an alternative that is strictly worse than another alternative you (must have) considered*. This is the essence of the *compliance dilemma* as presented in section 6.3.1. Whether to use the first or second conceptualization depends to a large extent on the knowledge representation language one uses: the first criterion is a classical one, while the second one is epistemic, and undoubtedly the better account of how norms are actually used.

Classification systems in LKBS will usually use the classical conceptualization (cf. generally section 4.6) – some behaviours are disallowed and the alternatives are irrelevant – with the risk of creating compliance dilemmas. For planning one would generally want to consider the second approach, which allows one to take into account compliance dilemmas and resolve them, but of course with the risk that alternatives are overlooked that others, that ascribe an intention to your behaviour, did not overlook.

⁹loosely based on [268]

On the logical level we also have to distinguish ordering either alternatives or sets of alternatives. In a normative theory that does not contain conflicts, this distinction is of less importance than it would be in a true logic of preference, because we are sure – given the definitions of deontic operators in terms of preference – that the preferences are all properly partitioned (if $P \succ Q$ then $|P| \cap |Q| = \emptyset$). Because we are assured of this, as long as we are able to determine in which order to consider application of conflicting normative statements these do not pose a complex problem, but one which will not be addressed in detail in this book.

Exposing the normative order intended by the legislator is however not merely a matter of envisioning and ordering the categories of cases that the legislator's rules distinguish, as sections 4.5 and 4.2 pointed out: normative order always exists, regardless of whether it is formalized, and the legislator often just tinkers with the details of an existing social order in society. Sometimes the actions of the legislator only confirm what is already *normal*, by formalizing it, or perform minor repairs in existing voluntary normative order, and sometimes the legislator only tries to create the right conditions for acting in one's own interest without obstructions or harming others (e.g. facilitating a free market). Institutional reality and the intended normative order are different things.

The legislator acts on an existing normative order, or more accurately on his preferred explanation of the existing normative order, and performs legislative actions with the intent to change it. The same is true on the level of the institution, as pointed out in the previous chapter: the sources of law are not a blueprint of the institution, but a log book of relevant legislative changes to it. The legislator relies on three important mechanisms to effect changes in the normative order: public announcement of the rules, the unique ability to effectively organize punishment, which underlies the normative rule, and the widely held preference for predictable social interactions among the addressees of his legislative acts (the scripts of section 4.5).

The legislator's acts obviously sometimes *fail* to achieve the intended effects on society. Besides that, the legislator is usually not a single natural person: it is highly unlikely that the actual people involved (for instance a parliament) share a coherent theory of an ideal normative order. The legislator only has intentions in the same sense that a company has a strategy: we are only likely to take the notion seriously if written evidence of the strategy is published in the name of the entity.

Institutional reality works separately from normative order, but some notions do appeal to normative order, or the intentions of the legislator: rights, compliance dilemmas, conflicts, unrealizability of norms, etc. An intended normative order is ascribed to the legislator, because we need to identify when the legislator fails to create the intended normative order. Section 6.7.2 will show that the courts go beyond the legal rules in the interest of creating a coherent normative order.

A subjectivist would argue that the way we motivate our actions is beyond rational criticism. This especially applies to the motives of the legislator, who is not bound to any requirements of rationality and not a rational agent in the normal sense. Suber (cf. [260]) expresses this view when he observes that the legislator can in principle abrogate any common sense knowledge or widely accepted logical inference and replace it with legal fiction. The legislator is obviously not omnipotent or omniscient, but correction

of the legislator should be achieved through the democratic process and not the legal system.

Action can however still be rightly criticized as irrational. The first kind of mistake is when action is based on false beliefs. If one accepts that the action a I intend to do in actual situation s and the action I would intend to do in the same hypothetical situation s are the same, which is widely accepted, then I should be willing to change my behaviour if false beliefs I hold are corrected by someone else. If I learn that I am actually in situation s' and I already believed that in a hypothetical situation s' I would intend to do a' instead of a , then I will intend to do a' instead of a . Actions are after all understood to be precisely those events that are causally explained by the intentions of the actor (confer section 4.2).

The second kind of mistake depends on the possibility of a difference between my revealed preferences and my beliefs about the preferences that motivate my behaviour. Section 2.4.1 explained the silliness of equating revealed preferences and motivating preferences in the form of a joke: revealed preferences are ascribed, while motivating preferences are policies or *rules* one adopts. If there are such rules, then it is possible to correct them by reflecting on the effects of one's actions.

As noted in section 6.5), a preference theory represents the *norm* implied by one normative rule, or if preferred a set of normative rules provided they are consistent with each other. A set of preference theories, and a preference theory about choosing between them, describes a normative order. Normative rules are interpreted as subjunctive propositions about *betterness*, made by a legislator, that are intended to play a role as motivating preferences in decisions.

To explain the difference between these two views on preferences, we have to delve a bit deeper into the notion of choice. An *induced decision rule* $\mathcal{C}(\mathcal{A})$ for a menu \mathcal{A} selects an alternative from the menu. Decisions can be conceived of as applying an induced decision rule. A rule $\mathcal{C}_\succ(\mathcal{A})$ can be directly induced from a preference relation \succ if it is completely preorder (complete and transitive) and asymmetric:

Definition 1. (Induced Choice Rule) Given a relation \succ , the Induced Choice Rule \mathcal{C}_\succ is defined by: $\mathcal{C}_\succ(\mathcal{A}) = \{x \in \mathcal{A} : x \succ y \text{ for all } y \in \mathcal{A} \text{ and } y \neq x\}$

That a preference relation is transitive and asymmetric is a reasonable assumption to make of a rational agent. There are however two different ways of looking at the relation between preference and choice. One theory, let's call it the *revealed preference* theory, postulates that preferences are revealed by choice between alternatives. That is:

Revealed preference : $a_1 \succ a_2$ if the agent chooses alternative a_1 over alternative a_2 .

That choice always entails a conscious comparison of *all* alternatives and therefore a preference is not entirely obvious; people enthusiastically choose between equally valuable products on a daily basis, and for instance show no signs of distress when confronted with the choice between a well known and an unknown alternative in the supermarket (cf. [138] for a discussion). Revealed preference theory solves this problem by postulating that if a comparison would result in preferring a_1 over a_2 then I already

prefer a_1 over a_2 but didn't *know* it yet. Our preferences can be discovered through comparisons. A corollary of this thesis is that preference orderings over alternatives are always complete.

The other theory, let's call it the *motivating preference* theory, postulates that we can be truly indifferent to alternatives while still making choices. That is:

Motivating preference : the agent chooses alternative a_1 over alternative a_2 if $a_1 \succ a_2$.

An incomplete preference ordering cannot be used directly to induce a decision rule. Instead we can only say that a certain decision rule is consistent with a *preference theory* (cf. section 6.5), or is a *bisimulation* of it. It is for instance possible to induce a choice rule based on the assumption that any choice between alternatives to which the original choice rule is indifferent will do.

It would seem odd to distinguish *motivating indifference* – revealed indifference certainly doesn't exist – but such a notion is useful in law. It is obviously meaningful if one adopts motivating preference *and chooses* between incompatible incomplete decision rules like one does with (legal) choice rules because one is not able to formulate a single coherent one. Permissions play the role of a motivating indifference, but *only* in the sense of restricting how we can extend the normative order by formulating additional normative rules. The addressee is only bound by \succ , not by \succeq . As long as a legislator values freedom of choice, the legal system will always formulate incomplete choice rules.

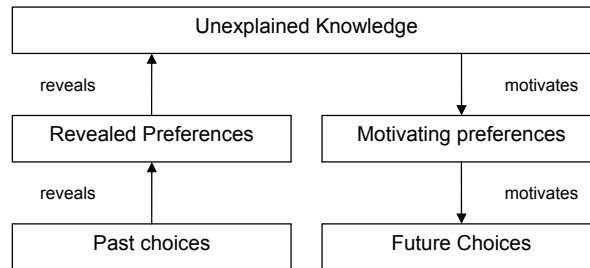


Figure 6.1. The problem with the idea that the preferences are revealed by choices, and the choices motivated by preferences is that these two categories of preferences cannot refer to the same thing.

Revealed and motivating preferences are clearly different (see figure 6.1): we can in fact speak of a legislative monitoring cycle: to legislate involves reflecting on choices one saw people making in the past, formulating hypotheses on why these choices fail to bring about the intended normative order, and formalizing new rules to guide choices in the future.

One can of course criticize the coherence of the normative order as created by the legislator, or point out that this and that is against some important value. The idea that one can simplify the system by isolating core values from it is however simplistic. Attention for undesirable outcomes in specific cases tends to increase the complexity of the law.

To understand why, we need to appeal to the property of preference independence (cf. for instance [240]). Hansson ([138]) distinguishes between preferences among incompatible alternatives, which he refers to as *exclusionary preferences*, and preferences among compatible alternatives, which he refers to as *combinative preferences*. When I choose a red car over a non-red car, I am exhibiting an exclusionary preference, because the alternatives are incompatible. My preference for having a cheap car rather than a red car, by contrast, is combinative, because the alternatives are compatible. The norms in legislation are all examples of a particular type of statements of exclusionary preference: $|\alpha| < |\neg\alpha|$ and $|\alpha| \cap |\neg\alpha| = \emptyset$. If these are preferentially independent, they are easily composed into a coherent system. If not, we will be faced with the situation of having to choose between them on occasion.

But in law the property of *preference independence* clearly often does not hold¹⁰. Given propositions P and Q , it is not generally the case that Q is preferentially independent from P . What is understood by preference independence is easiest explained by an example: Let P mean *running over a pedestrian in your car* and Q mean *calling 911 and telling them you ran over a pedestrian*: observe that it is entirely possible and reasonable that P is disallowed, and Q is allowed if P and disallowed if $\neg P$. This is essentially what happens in the contrary-to-duty situations discussed in section 6.4.

This strictly speaking rules out the use of many valuation tools used in decision theory for choice between alternatives like additive value functions, multiplicative utility functions, and other economic concepts based on the assumption of preference independence (cf. [169, 240]). Decision making subject to legal constraints therefore does not lend itself very well for approaches based on the *expected utility hypothesis*. The use of decision support methods like Multi-attribute Utility Theory (MAUT; cf. [104]) is ruled out by this criterion.

This lack of independence between evaluation dimensions makes it hard to conceive of underlying values in law that are not just defeasible-motivating-preferences-in-a-context, just like the original legal norms found in normative rules. Specific norms may sometimes affirm more general ones – the notion of *affirmation of an imperative* (cf. 6.3) – but legal norms, regardless of whether they are dressed up as fundamental principles or values, are always defeasible if we interpret them as a coherent normative order. The intentions of the legislator cannot be reduced to a number of independent core values that happen to be expressed in context-specific settings in the form of a specific norm.

Still, this does not discourage anyone from ascribing underlying preferences (not just intentions) to the legislator based on his observed actions, to analyze the normative-order-creating actions of the legislator using another, smaller, set of norms – norms of analysis – or for courts to attribute mistakes to the legislator and to fill in the gaps left in the implied normative order over time.

As alluded to in section 4.5.2, preferences are generally attributed to *agent* roles in a script; Human beings cannot be attributed a single comprehensive agent role, and the total set of preferences ascribed to them in order to explain their behaviour need not be consistent. For normative rules the same is true: normative rules apply to an

¹⁰I am not claiming that it does hold in some other domain, just that in law it *clearly* doesn't.

agent role.

6.7.1. Preferences and Scripts

When an agent makes a decision what to do, he has to deal with uncertainty. He has to weigh the utility of the foreseen possible outcomes of the alternative actions he can perform with their likelihood. This is the standard economic formulation of a decision problem, as solved by the expected utility hypothesis (see section 2.4.1).

The most important wild card an agent has to deal with in social interaction is the behaviour of other agents. The outcome of his actions are usually most sensitive to the decisions of the other involved agents: the preferences they have determine the epistemic preferences he should have towards recognizing their intentions. Even in a deterministic planning environment like chess this is shown to be very hard, while chess has a relatively low degree of complexity compared to potential social interaction between agents.

In real world planning a relaxed version of the problem is solved: because other agents available as resources in the planning problem are clearly not obviously in direct competition with the problem solver, the planner is limited to executing a number of social *scripts* that usually work: the alternative outcomes of an action are reduced to the one that normally occurs. This is considerably closer to what humans do, as discussed in the reflection on normal behaviour in section 4.5.

In the Hohfeldian scenario of section 4.8.2¹¹ this means that two expected scenarios are open for *a* who offers something for sale to *b*:

$$\begin{array}{c} \text{OFFERFORSALE}_a / \neg \text{OFFERACCEPTED}_b \\ \text{OFFERFORSALE}_a / \text{OFFERACCEPTED}_b / \text{SUPPLY SOLD ITEM}_a / \text{PAY}_b \end{array}$$

There is a norm n_1 which tells us:

$$\begin{array}{c} | \text{OFFERFORSALE}_a / \text{OFFERACCEPTED}_b / \text{SUPPLY SOLD ITEM}_a | \succ_{n_1} \\ | \text{OFFERFORSALE}_a / \text{OFFERACCEPTED}_b / \neg \text{SUPPLY SOLD ITEM}_a | \end{array}$$

There is a norm n_2 which tells us:

$$\begin{array}{c} | \text{OFFERFORSALE}_a / \text{OFFERACCEPTED}_b / \text{SUPPLY SOLD ITEM}_a / \text{PAY}_b | \succ_{n_2} \\ | \text{OFFERFORSALE}_a / \text{OFFERACCEPTED}_b / \text{SUPPLY SOLD ITEM}_a / \neg \text{PAY}_b | \end{array}$$

Agent *a* prefers compliance with n_1 and expects and prefers compliance of *b* with n_2 . Agent *a* moreover prefers the first scenario over the second, while agent *b*'s choice is unpredictable.

¹¹ Assuming a particular order here, and simplifying the representation. Since these are institutional actions, a real world action of paying may for instance constitute accepting the offer. In addition the *deadline* problem of section 4.7.2 is left unaddressed here.

In the case of normative rules this observation for instance leads to the prediction that people generally comply with the norms, but this doesn't tell us what they will do with the alternatives they have for changing or retaining their legal position. What legal positions people generally prefer to have is a separate modeling problem: as stated in section 6.7 the quality of a legal position is reducible to the normative rules, but in the end the question is whether b is interested in $\text{SUPPLYSOLDITEM}_a/\text{PAY}_b$ over $\neg\text{PAY}_b$: if b is, then b improves his position by OFFERACCEPTED_b . In the case of a sale that is hard to predict, since it depends on the item in question.

This is certainly not always the case: people generally like rights, eligibility to benefits, powers, privileges, immunities, and dislike duties, disabilities, liabilities, etc. Although planning is often goal-directed, there is a strong positional aspect (cf. [167]) to it, certainly when positional improvements come cheap. In predictable settings like the one described in section 5.2.1 (cf. [272]) this kind of thing goes without saying: very often we can go a long way by assuming that other people have the same preferences as we do.

Whether establishing a legal fact is a good or bad thing for an agent is generally clear; The legislator leaves clear hints how a concept is to be evaluated by the words he chooses, certainly relative to a setting, and if these didn't already have a clear evaluative interpretation, they will acquire one over time:

$$| \text{TAKING} \sqcap \text{THEFT} | \prec_p | \text{TAKING} \sqcap \neg\text{THEFT} |$$

A normative rule whose case is expressed by one concept, *makes* this concept value-negative or value-positive (cf. [138, 139]). In many other rules concepts also become so over time because of the constellation of rules around them. A legislator would not generally change this evaluative meaning a concept has acquired without inventing a new word to describe it¹².

In relation to scripts, section 4.5.2 distinguished between the physical human being and the *agent* role he plays in a script. The person plays an agent role, the agent fills the *actor* role, and other roles, in an action. The agent is identified by the set of actions that is performed in that role, but in addition certain preferences relevant to his role can be ascribed to him (ascription being the application of a default role when it is consistent to do so). It is the agent, for instance a buyer or seller, with the actions that characterize his role, and the default set of preferences ascribed to him, that explains the small set of normal executions of a script.

The preference theory attributed to the agent can be abstracted to a reified decision rule believed in by the (player of the) agent, just like the rule represented in the source of law, although this a mental concept and not an institutional one: \prec_a relation is a relation between domain $| \exists \text{disallowedBy}.\{a\} |$ and range $| \exists \text{allowedBy}.\{a\} |$, that describes the decisions normally made by agent role a , and works like the normative betterness relation in section 6.7.

Some actions, like the action of buying/selling, involve multiple actors. A completely passive role in the action is filled by the *patient*, who may also have preferences

¹²This is precisely the issue around political correctness, which is itself pejorative (i.e. value-negative): choice of words matters.

attributed to him, but doesn't take a decision to play the passive role: the patient of a murder for instance generally prefers not to be murdered. For planning this is less relevant, but for the legislator this can play a justifying role for certain normative rules. Note that an agent does not need to have explicit expectations about the preferences of others: these are already apparent from the agent role attributed to them.

The action, agent role, and rule are reified and described by OWL concepts. The *script* as intended by Schank (cf. [246]) is a theory, a set of logical propositions connecting action, the agent roles involved in their performance, and the preferences attributed to those agent roles. The process of recognizing the execution of a script is however an ampliative activity, and involves default rules. In the context of law this was explained in terms of constitutiveness.

Some questions remain open after consultation of the relevant literature (here i.a. [138, 241, 267, 100, 102, 101]) on preferences: the nature of the relation between preferences of persons and those of agents, and the nature of the relation between agents/persons and preferences. No name is proposed for the property that links preferences and agents/persons anywhere: the agents simply *have* them. In addition, not everyone will agree that the script attributes *preferences* to the agent: in the context of a specific execution of a script, it seems more appropriate to speak of plans, of goals, of reasons, of intentions, etc. In my view these are more specific roles or epistemic derivatives of the preference, and to fix the operational semantics of these terms we would have to commit to a more specific theory of planning than the one proposed in section 4.3.

As Girard eloquently pointed out (cf. [176] about the legacy of the French historian and philosopher, and [118] for a discussion of empirical evidence for *mimetic theory*), *imitation* is a driving force behind human behaviour: we can therefore expect to share a major part of our preferences because we acquired them from each other. Presumably this is why we can discuss behaviour in terms of prototypical scripts, and recognize the intentions of others.

One potential misunderstanding should however be cleared up. Holding the same preferences is hardly a reason to be natural allies: competition for scarce resources is in fact most easily explained by the fact that people have the *same* preferences. People compete for the same things because they value them equally, and usually for exactly the same reasons. When two drivers collide at an intersection, they do so because they *both* generally prefer not stopping over stopping. When a driver and co-driver disagree on what is the appropriate speed to drive on the highway, they would seem to exhibit different preferences. Whether we have the same or different preferences is not a guide to whether we are likely to get into conflict. Having the same preferences does not make us cooperative: some things are simply better achieved by cooperating and others better by competition.

More relevant is the distinction between I-preferences and we-preferences, akin to the I-intentions and we-intentions distinction briefly mentioned in section 4.2 (which apply to specific actions, and therefore the actor in them, instead of to agents). In [243] we find the same distinction as private preferences vs. communal preferences. The I/we distinction applies to the object of the preference – the agent whose action is evaluated – and not to the holder of (believer in) the preference: I don't consider

whether “groups” can be said to have a preference; What matters is whether they can assume an agent role.

In case of scarcity it is more realistic to have an I-preference *I prefer that I take the best seat* over a we-preference *I prefer that we take the best seat*, for instance. In cases where we are all better off by acting with the same intention the we-preference *I prefer that we all yield for traffic from the right* will generally win from the I preference *I prefer that I have the right of way* since it is clearly more reasonable.

In [243] some additional interesting classifications are found that determine whether social coordination and publicly announced rules are likely to arise or not. [243] also points out that communal preferences in themselves are also not necessarily cooperative in nature.

Sharing the same set of preferences, regardless of whether they are we-preferences or I-preferences, can however go a long way in practical planning. The representation of we-preferences of course doesn’t present us with any problems: these are simply the norms or preferences attached to roles in certain actions, and the preferences relate to the action executed in that role.

These can of course be restricted to specific populations as the players of roles: one might for instance have certain scripts in mind regarding purchases over the Internet, but I may distinguish between buying from Dutch web shops – for which I have more elaborate scripts about the way faulty products, or products that don’t meet my expectations, are handled – and buying from other web shops¹³.

Private preferences, or *subjective norms*, are default rules indexed to an individual, for instance: (DEFAULT r (known.BUYER $\sqcap \{i_{AlexanderBoer}\}$) ...). These individuals may be the *self*, other actually existing ones, or prototypical ones¹⁴. The case for attaching a special status to the self, with a special i_{self} constant, is interesting (cf. generally [181]), and sometimes pragmatically necessary, although such an identification is of course rather context-specific for a knowledge representation.

How the notion of a script is worked out in a knowledge representation, based on the represented legal normative rules, and subjective and communal normative rules of agents, and background knowledge on actions, in practice often depends on the context of use. In an administrative setting the legal rules will be seen as a set of constraints on a *business process* to be designed. For a planning system scripts are often interpreted as plan operators, and for a design system they are interpreted as design operators. Any attempt at standardization in this direction (e.g. *STRIPS*, *ADL*, *PDDL*, cf. generally [240]) is inextricably linked to a certain type of planning or design algorithm. In case law they are the thing an interesting set of related case positions deviate from ([79, 80, 9] for instance discuss findings with the example script of chasing wild animals – hunting game, fishing – and case positions involving interference with it). In comparative law we may also see the case positions, or interests of stakeholders that are being protected by legislation.

¹³Only 16% of the Dutch, and 7% of Europeans, bought a product abroad in the last year; 62% of the Dutch bought a product from a Dutch web shop (*Press release of EU commissioner Kuneva of consumer affairs, June 20, 2008*). A compelling explanation is unfamiliarity with the exercise of consumer rights abroad.

¹⁴See *prototype* in section 3.4.3.

The various representations encountered are motivated more by pragmatic constraints than by a coherent view of scripts as mental objects. This work does not significantly improve on alternative representations by proposing a representation of these preferences. The preference itself can also be rightly criticized for being cognitively not very plausible. The relation between legal rules and the scripts they invoke is however central to understanding the positional evaluation in terms of for instance Hohfeldian categories, adjudication of cases and use of case law, discussed in section 6.7.2, and comparative law, discussed in section 6.7.3.

6.7.2. Case Law

When we think of case law, we generally think of court cases that involve a serious conflict between a plaintiff and a defendant, in most cases decided either for the plaintiff if some future or past behaviour of the defendant is determined to be disallowed, or for the defendant if the behaviour was not determined to be disallowed. What legal consequences this decision has depends on the type of legal procedure (tort, criminal, administrative, etc). Section 4.8 pointed out that the disallowed-decision of the court normally does nothing more than granting a potestative right to the plaintiff or a third party, generally to do something unpleasant to the defendant.

When considering the daily activities of the court, it is however important to make a distinction between the conflicts between plaintiff and defendant that are of interest here, and notarial activities, where a judge acts as a witness in a transaction between two parties that cannot be characterized as plaintiff and defendant in order to establish legal facts. Even the vast majority of conflicts between private parties handled by the courts are routine cases that can be characterized as simply denying or granting potestative right to a plaintiff: center stage takes the question whether plaintiff can make the court establish a certain legal fact that is advantageous to him.

Obviously this type of case is the most obvious candidate for extra-judicial solutions, like LKBS and mediation: the current trend is to try to move these cases out of the hands of the courts. This is not necessarily as easy as it may appear. The defendant can refuse to cooperate with any extrajudicial procedures if the potestative right to go to court remains in existence, and may do so on purpose to intimidate the plaintiff and delay the resolution of the issue in favour of the plaintiff. The plaintiff may also file a lawsuit with the mere purpose of intimidating the defendant. These scenarios are obviously not likely to make it into case law.

The activity of courts of law can be described in terms of two important principles from a knowledge engineering perspective: Firstly, the courts apply the applicable legal rules to the cases brought to them, and they are themselves bound by legal rules. Secondly, the courts are expected to be predictable, even in the absence of relevant legislation, or if the relevant legislation is ambiguous: if two cases are substantially the same, the courts should come to the same decision in both cases.

Stare decisis, Latin for ‘to stand by that which is decided’, is the general maxim used in common law that when an issue has been settled by a court decision, it forms a precedent which is not afterwards to be deviated from by other courts. The doctrine of *stare decisis* does not completely prevent re-examining and overruling prior decisions,

but it is considered (cf. generally [260]):

”a fundamental jurisprudential policy that prior applicable precedent usually must be followed even though the case, if considered anew, might be decided differently by the current justices. This policy [...] is based on the assumption that certainty, predictability and stability in the law are the major objectives of the legal system; i.e. that parties should be able to regulate their conduct and enter into relationships with reasonable assurance of the governing rules of law.” (Moradi-Shalal v. Fireman’s Fund Ins. Companies (1988) 46 Cal.3d 287, 296.)

A party arguing overruling a precedent faces a difficult task, roughly in proportion to the age of the precedent, the extent of society’s reliance on it, and its conceptual and functional similarity to other related rules of law.

A very similar notion in civil law is found in *jurisprudence constante*, which is basically a weakened formulation of stare decisis: the longer a series of identical decisions in identical cases based on the same legal rule by the courts becomes, the more determinative the decision found in the series becomes for future cases. A conceptual difference between stare decisis and *jurisprudence constante* is that *jurisprudence constante* denies the possibility of courts deciding a case based on precedent alone, without appeal to a legal rule found in legislation, and denies that the court decision potentially creates a rule.

The difference is mainly a justification issue: the precedent case can be cited as an example of why the used interpretation of a legal rule is the correct one, but citing a precedent case does not fulfill the burden of proof of backing one’s decision with a legal rule from a formal source of law (a constraint along the following lines: \exists representedBy. \exists metalex:resultOf. \exists agent.FORMALLEGISLATOR).

It is not hard to see how social mechanisms reinforce the stare decisis/*jurisprudence constante* doctrine: 1) Courts naturally don’t like to be overruled by higher courts, 2) legal professionals are naturally interested in discovering patterns in what the courts do, and in anticipating future court decisions with these patterns, and 3) society at large likes the application of the law to be predictable. The design rationale for legal rules (see section 4.1) is itself based on the stare decisis argument.

One could claim that the stare decisis notion has somewhat tautological character: cases to be decided obviously never are the same; That two cases are the same is as much a result of the choice to come to the same decision as the other way around. If one wants to come to a different decision, doesn’t one simply point to a difference between the present case and the previous one?

But things aren’t that simple: new facts of the old case can no longer be established to compare them to the new case, whereas determining new facts of the new case is still possible. It is not possible to determine whether the new case deviates from the old case by considering an additional fact that is not in the *case description* of the precedent case. It is therefore the case description that binds future court decisions, and not the case. The link between case description and decision is a categorical one, and the precedent case therefore definitely postulates a *rule*, regardless of whether we are allowed to call it a *legal rule*.

It is perfectly reasonable to represent it as such, and in exactly the same way as other, legal, rules. We do however have to keep in mind that the rule postulated is the

totality of the mapping from case description to outcome: there is no doctrine *stare dictis* (to stand by or keep to what was said) or *stare rationibus decidendi* (to keep to the rationes decidendi – the justification – of past cases). A court decision postulates just one rule: the immediate generalization of the case and its outcome.

A useful framework for relating legal rules expressed by legislation to rules in case law is the framework for case law theory construction developed in [18, 19, 79, 80, 9]. In these publications, Trevor Bench-Capon, Giovanni Sartor, and others, introduce a framework for relating case law, which describes case law theory in terms of 1) cases and the outcomes of these cases, 2) factors in terms of which the cases are described, 3) abductively formulated rules linking sets of factors to (postdicted outcomes), and 4) values promoted or demoted by having the presence of a factor weigh in favour or against deciding for the plaintiff. A theory is understood to comprise:

1. a selection of relevant cases from the case background;
2. a selection of factors from the factor background;
3. a set of rules linking sets of factors to outcomes;
4. a set of preferences amongst rules; and
5. and a set of preferences over values promoted by the factors within the rules.

Figure 6.2 explains how cases, factors, and values relate to each other. Note that I use the vocabulary of *revealing* and *motivating* from sections 6.7 and 2.4 here to relate the layers. In the original figure in [9] a distinction is made between *explaining* the case background, and *determining* the new cases. On an operational semantic level both relations between factor-set preferences and case decisions are of the same kind: the preferences between sets of factors motivate the choices, and the background of choices from the past reveals the preferences between sets of factors.

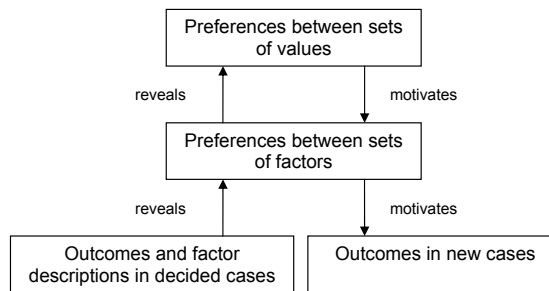


Figure 6.2. Construction and use of case law theory, adapted from [9].

The theory is developed in [18, 19], backed by the results reported in [79, 80], and finally consolidated in [9]. In [80] a tool, named AGATHA (Argument Agent for Theory Automation), is described that generates case law theories from a seed case and a body of relevant precedent cases. It is encouraging that AGATHA in a well-understood domain produces better theories than earlier handcrafted theories (viz. [79]; cf. [80]).

The *factors* are the propositions of interest that indicate a relevant brute or institutional fact, the values are associated with norms of analysis in this work, and the rules are the legal rules we find in legislation.

The outcome is – as is appropriate for tort law – phrased in terms of *deciding for the plaintiff* or *deciding for the defendant*. This is a generalization of the decision alternatives: one of the possibilities is that the case brought to the court is unlawful if decided for the plaintiff. The vocabulary – deciding for plaintiff or defendant – applies equally to criminal law, but seems out of place when the plaintiff is a public prosecutor. The rules linking sets of factors – to be read conjunctively – to outcomes in favour of the plaintiff or defendant, are simply datalog rules linking a *generic case* in terms of [266] to an outcome, for instance one of *allowed*, *disallowed*, *silent*, in which case they would be easily recast as normative rules. Bench-Capons *rule* is the same type of entity, with the same characteristics, as the indicative legal rule in this work. To decide for the plaintiff is to establish the legal fact he proposed.

Factor background information is represented as tuples (f, o, v) , where f is a factor, outcome-favoured o is one of {PLAINTIFF, DEFENDANT}, and v is the value promoted by that outcome. An outcome favouring the plaintiff usually maps to *disallowed* if the main issue to be settled is normative, while an outcome favouring the defendant should map to either *allowed* or *silent* if the rules that the court is supposed to apply do not apply to the case, or the burden of proof for them is not met.

The distinction between the approval or the silence of the law as we find it in [266, 189, 172] is not readily apparent in case law: one could argue that there is a difference in cases decided for the defendant between court decisions that posit positive *allowed* rules to decide for the defendant that have an influence on later decisions and *silent* court decisions that merely note that any grounds for deciding for the plaintiff are simply absent. There are good explanations of why the possibility of a *silent* court decision attracts less attention in private law than in areas like criminal law, where potential murderers sometimes walk out of court because of the law's silence:

1. since the conflict is between two private parties, and involves considerable investment and risk, it is reasonable to expect that the conflict only ended up in court because the involved parties genuinely *disagree* on the outcome, which suggests that there are at least some rules favouring the plaintiff;
2. court decisions that are *silent* rarely have an effect on jurisprudence and are generally less likely to end up in case law studied by legal professionals; and
3. there is no good rule for distinguishing silent decisions from allowed decisions, both against the plaintiff.

Bench-Capon prefers to relate values to single factors only, and not to sets of factors. This choice appears to be motivated by an *ontological commitment*, being that factors are assumed to be naturally *evaluative* concepts, inextricably linked to a certain value (cf. [235]). Factors are similar in function to the relevant *contrasts* of psychological geometry introduced in 6.1.1: the axis of reference is the norm or the value. The evaluative reading applies to any fact, not only those involved in deciding whether something is allowed or disallowed. Any factor that is cited in favour of one side argues that there is a reason to prefer the case of that side, i.e. the conclusion

it is arguing for. In other words, if $\alpha \prec_a \neg\alpha$ is a motivating preference in a decision, then α is a relevant factor.

This choice affects the granularity of modeling. In traffic, for instance, we can comfortably link a legally relevant phrase like *defendant was making a special manoeuvre* to a value in the right context, but not a more neutral one that doesn't suggest evaluative context like *defendant was making a left turn*. By doing this, the conceptualization does not take full advantage of things like *exceptions* on the level of factors: the presence of one factor could be entailed by the presence of another factor that promotes another value and points in the direction of an opposite outcome. The relation between both factors can of course be found by induction from the cases, but the information that there are logical subsumption relationships between the factors is not part of the case law theory. The framework on the other hand does not prohibit expansion with a logical level at a lower level of granularity to explain subsumption relations or possible overlap between the factors, just like the separation of the allowed and disallowed case of a normative rule is only a start to more complicated modeling of the normatively qualified behaviours.

The courts have some freedom in filling in the gaps left open by the legislator with the rules they create. Clearly the courts have the freedom to apply their own theories of brute reality to a case: the legislator does not supply such a theory. One could therefore say that they are free to create new elaborations of indicative rules and requirements, which, if they have an effect as precedent, will function in the future as affirmative rules affirming a more general rule in formal law (cf. *affirmation* in section 6.3).

They however also have a limited power to change institutional reality: since the recognition of conflict and the solution of conflicts by way of *lex specialis* also depend on one's theory of brute reality – it is only relative to such a theory that one can decide that one constellation of facts is more specific than another one (cf. section 6.3.1) – the court's decisions, if they have an effect as precedent, even create new *choice rules* (cf. section 6.6.3). The court can obtain the freedom to *not* apply a legal rule by constructing a case for a conflict or dilemma that gives him the freedom to choose between that rule and another one.

What the courts presumably do when they use this space left open by the legislator – at least when they are not being “activists” – is simply connecting the dots: they ascribe intentions to the legislator, and to the previous courts that created important case law, and try to infer the intended normative order and decide in accordance with it.

The courts in this way become a motivator for legislative change. New legislation does not come out of a vacuum: there are reasons to draft new legislation, and one of these reasons is to periodically *clean up* legislation¹⁵ by codifying developments in case law and other legislation, for instance from a higher legislator (the European Commission to name one), that have come to be considered *opinio iuris*. The need to do so is obviously greater in civil law jurisdictions than in common law jurisdictions, where case law is permitted to take the center stage.

¹⁵In the Netherlands this is generally called a *veegwet* (wipe act), literally intended to dust off an old act by codifying relevant *opinio iuris* into it.

If the legislator disagrees with the intentions ascribed to him, he must explicitly deny the existence of rules as articulated in case law, which will lead to increasing elaboration of legal rules over time. To block application of undesirable jurisprudence rules, the legislator has to define new *requirements* or *applicability rules*. There is no obvious formal method for the legislator to repeal case law, although the intent to block the application of certain case law can be made explicit in legislation or legislative documents accompanying and explaining the intent of legislation. MetaLex at present doesn't cover this eventuality.

Case law can be largely treated in analogy with formal legislation, including the application of the choice principles of section 6.6, but the analogy breaks down when we consider versioning. The legislator usually takes care to restrict the applicability of rules to a specific time interval. Case law can become irrelevant over time because it was 1) overruled by newer and superior case law, 2) rendered inapplicable by additional requirements or applicability rules in formal law, or 3) because it was absorbed into formal law.

6.7.3. Legislating and Comparative Law

A last context to be discussed in this book where intended normative order in relation to scripts plays an important role, is the activity of comparing alternative methods for creating an intended normative order itself.

Comparative argumentation about the law is the province of the academic field of comparative law, but it is important to realize that comparative arguments about legislation are not only made by academics. These arguments are also commonplace in macroeconomics, in politics, in legislative drafting, and finally in the court room in the choice between individual legal rules. It is this use of comparative argument that is relevant to legal knowledge engineering. Comparative arguments are made wherever choices between rules or sets of rules are made, but often they are not recognized as such.

There are three common types of comparison of legislation (cf. [42, 50]):

Comparison of alternatives Proposals for legislation addressing the same problem are compared to judge which one is better according to preconceived norms of analysis;

Vertical comparison in time Versions of the same legislation in time are compared to determine the effects (costs and benefits) of changes of legislation on behaviour, products, etc;

Horizontal comparison between jurisdictions Legislation addressing 'similar' things in different jurisdictions is compared to inform others about the effects (costs and benefits) of moving themselves, their property, products, or services over the borders of a jurisdiction.

The legislative design problem, at least the part that involves constructing and choosing the preferred candidate solution, is itself a reformulation of a comparative law problem.

Montesquieu is regarded as the father of comparative law. His comparative approach is obvious in the following excerpt from chapter XI (entitled *How to compare two*

different Systems of Laws) of book XXIX of *De l'esprit des lois* (viz. [91]):

To determine which of those systems [i.e. the French and English systems] is most agreeable to reason, we must take them each as a whole and compare them in their entirety.

Comparative research in Law usually claims to compare coherent legal *regimes*, including:

- The aspects of social life that law proposes to regulate;
- The extent to which law actually regulates social conduct;
- The attitudes about the proper structure and operation of a legal system; and
- The dedicated legal organizational structures, agents, and processes that exist for legislating, execution of the law, and adjudication.

When we however look at the practical *functions* of comparisons (cf. for instance [174]) – when they are not academic – then we see that different variants of sources of law are indeed the candidates considered and chosen between. A harmonization effort in law, for instance, usually involves a number of comparisons between regimes, with different purposes: They are compared to find a new *common core*; To separate the good from the bad legislation; To repair the latter. Relevant external legislation is involved in legislative drafting to establish its suitability for a *legal transplant* to another jurisdiction. Surely all of these things would not be possible or useful if the legal regime does not in the end boil down to the sources of law: the rules are the main objects being harmonized or transplanted. The observation that a thing that works in one jurisdiction is not at all guaranteed to work in another one is undoubtedly true, but of limited usefulness.

Besides that the legal regime also remains largely the same in some scenarios. If one compares alternative legislative changes to a new situation, the old situation that is being changed is always the same. A migratory regime is designed and the costs of migration from the old to the new situation are also evaluated, to take into account the costs of migration, and so that the individual and collective costs of migration can be minimized and fairly distributed.

It is realistic to compare sources of law of similar societies, certainly in areas like taxation where uniformity of approach has had a clear and obvious economic value for a long time, and a shared vocabulary already *de facto* exists (cf. for instance [106]).

In the last decades the interest in the problem of comparing and harmonizing legislation has been steadily increasing. One reason is the increasing legal convergence between governments in the European Union, and the increasing traffic of people over borders of jurisdictions. This leads to an increased need of administrations to know and understand legislation of friendly governments to be able to assist citizens and reduce negative consequences of movement.

Another reason is the increasing globalization of companies; Products and services are offered in many jurisdictions at the same time, and the product or service has to meet the provisions of all jurisdictions in which it is offered. Different regulations can also lead to differences in competitiveness for the product or service. For a financial product, for instance, it is considered important to qualify for tax deductions that make the product more attractive. For a medical product it is important to know whether it can be sold over the counter without a prescription.

At the same time the Internet also goes through a revolutionary growth phase, and there is an increasing interest in the use of information systems to make the problems created by increasing legal convergence and economic globalization manageable (cf. generally [42]).

The discovery of the requirements of comparing legal systems and sources of law has a direct relevance to theory in legal knowledge engineering: it has the potential to create a more principled, more robust, and language-independent methodology for modeling legislation, because it decouples the model of legislation from the purpose for which it was modeled and makes assumptions about the surrounding context, and knowledge expected from users, explicit. Modeling only one set of regulations for a specific customer often fails to achieve this level of context-independence.

Many recent initiatives have recognized the problem of mapping legal vocabularies to each other and proposed standard ontologies of international legal vocabulary (see [42] and generally [112, 22] for recent references in computer science, or for instance UNIDROIT¹⁶ and Common Core; cf. [70]). The approach of integrating two ontologies representing two different versions of a source of law has been tried by [115, 116], under the heading of *norms dynamics*. Most projects start from the assumption that a taxonomy of legal concepts is the obvious way to explain the relations between concepts from multiple jurisdictions. The problem of comparing legal systems is thus reduced to an *ontology integration* problem; Multiple jurisdiction-specific ontologies must be integrated into one international ontology. General approaches and tools for integration of ontologies are discussed in for instance [226, 173, 210, 73].

There are several significant problems with this approach, which limits its usefulness. Firstly, observe that this book only classifies a minor part of the information contained in the sources of law as *legal ontology*. The subsumption problem for legal rules is still an open problem, as section 6.6 concluded. Even the problem of legal translation poses significant problems (cf. [42] for a worked out example): institutional realities exist *sui generis*¹⁷, and one cannot just claim that *public body* in the UK translates to *bestuursorgaan* in the Netherlands, or the UK's *theft* to Dutch *diefstal* or German *diebstahl*. In the case of theft the ontological mismatch is especially clear: theft is an act of appropriation, not necessarily by taking, while *diefstal* and *diebstahl* are acts of taking with the intent to appropriate, but a successful appropriation is not necessary.

One term may stand in for the other as a synonym in a certain context, but this doesn't make them the same concept¹⁸, as section 2.3 explained.

A recent comparative law view on why legal concepts like *theft* and *diefstal* may be considered to be sufficiently similar to function as a target for comparison is found in [273].

Most importantly, however, the ontology integration view does not account for the purpose of the comparison in a satisfactory way, and that explains why publications

¹⁶<http://www.unidroit.org>

¹⁷In a category of their own; As the only instance of their class.

¹⁸Look for instance into 7246/2003 – C5-0165/2003 – 2003/0811(CNS) for a total of 23 proposed amendments by the Hellenic Republic to clarify the meaning of the phrase *same criminal offense* in the context of application of *ne bis in idem*.

in the field of comparative law do not seem to be related to ontology integration at all. If comparative argument appears related to ontology integration, it is merely because the ontology happens to describe the relevant *factors* (cf. section 6.7.2) or *evaluative concepts*.

Comparative arguments are based on rhetorical comparison between two targets with respect to an axis of reference (cf. section 6.1.1), with the aim of uncovering interesting similarities and differences. This axis of reference is a *norm of analysis* (cf. [145, 146]), and is yet another norm which is based on the *subjunctive betterness* pattern, just like the legal norm, the communal norm, and the subjective norm.

To make clear what norms of analysis are, and how they are used, I will give some informal examples from the field of taxation, in particular taxation of capital income. Capital income can be taxed based on the actual (*ex post*) income data or the predicted (*ex ante*) income data. The Dutch *predicted capital yield tax* introduced in 2001 is an example of an *ex ante* tax on expected yield. Only a fictional risk free yield is taxed; additional gain or loss attributed to specific skills, labor, or information advantage is not taxed. The *capital gain tax* (where gain is taxed at the moment of realization of income) and the *capital growth tax* (where apparent capital gain is taxed at fixed intervals based on an estimate of what income *could* be realized from capital regardless of realization of income) are *ex post* taxes on capital gain. These tax risk free yield (the compensation for delayed consumption) as well as additional gain or loss attributed to specific skills, labor, or information advantage.

The following is a comparative argument based on a norm of analysis:

The consequence of the possibility of delaying capital gain taxation is that capital growth is taxed with varying rates depending on how long investors can afford to delay realization of income. This is inconsistent with horizontal equity.” (Cnossen & Bovenberg in [83])

The norm applied is known in the literature as the *horizontal equity* norm (those with equal ability to pay taxes should be assessed equal tax payments; cf. [252]). The comparatum is implicit: capital yield and capital growth are more equitable than capital gain tax. The author states that it is “inconsistent” with the principle, but that observation is only relevant because of the alternatives that *are* better implementations of horizontal equity.

Another example:

It is beyond dispute: the taxpayer could easily realize the predicted capital yield if he would invest his capital in risk-free investments. But what if he does not? Can we ascribe income to that taxpayer? An analogy from “human capital”: A taxpayer works 20 hours a week, and earns 30,000. The taxpayer could easily earn 60,000 for 40 hours a week. But what if he does not? Can we ascribe that income to that taxpayer and tax it?” (Vording in [83])

The comparatum is again missing: capital gain and capital growth tax are more effective than capital yield tax for some reason. The implicit norm of analysis is supplied by Cnossen & Bovenberg (also in [83]): “The principle of realization of income is based on the observation that tax should not have to be paid if the means are not actually available.”

The analogy with labour strengthens the argument in a slightly deceptive way: The difference between working 20 hours per week or *easily* working 40 hours a week

is hardly the same as the difference between choosing risk-free and risk-prone investments.

Average capital growth is closely related to the amount of risk associated with an investment. Not realizing the risk-free capital yield is usually a consequence of risk-prone investment strategies. Government may not wish to share in the fortunes of investors, or may not wish to *subsidize* risk-prone behaviour.

Additional such arguments are easy to construct: for instance, a fictional yield and an actual yield are easier to tax than an estimated one, because less information is required to determine it, a fictional yield results in a more predictable government budget than the other ones, etc. In all these examples the relevant difference is a rhetorical comparison relative to an axis of reference.

Rhetorical arguments for similarity were less easy to find in [83], but the field of law certainly asks for them, as the following examples show:

1. Art. 38(1)(c) of the *Statute of the International Court of Justice* directs the court to apply “the general principles of law recognized by civilized nations.” This not only entails discovering similarities between legal systems, but also confronts us with the problem of deciding which of its members the UN considers civilized nations and which not.
2. Art. 215 of the *Treaty of Rome* establishing the EEC provides that the non-contractual liability of the Community is to be governed by “the general principles common to the laws of the Member States.” Note that this formulation appears to make the intended similarities a moving target, dependent of when which member state acceded to membership of the EEC. The assumption is that the intended similarities are limited in number, and widely shared in the international community.

The question is who gets to decide whether things are sufficiently similar. In international law it is easier to define the norms that should be shared by member countries than to decide who is the authority that determines whether a member state complies with the defined norms. In harmonization and standardization the problem is which system moves to which other system. Harmonization and standardization are usually relative to another system that specifies common aims. In the EU we see this method in action: the commission defines the guidelines that member states must be consistent with, and it also defines the yardsticks used for determining whether a member state complies with the guidelines.

Interesting differences and interesting similarities are interesting because they indicate a positional difference or similarity in evaluation relative to a norm of analysis.

If the function of law is reduced to normative rules and the *allowed–disallowed* axis, then the problem of comparing two different normative orders would conceptually be relatively clear, although not necessarily easier to realize. In section 6.7 the idea of *envisioning* the normative order was introduced: what categories of cases does a normative system distinguish and how does it order them? The evaluative differences are found in pairs ordered in opposition by the two normative systems. This of course depends on being able to align the ontologies underlying the two normative systems, so that subsumption relations between categories of cases can be computed, and this

will be considerably easier for developments over time in the same jurisdiction, than between jurisdictions.

Institutional ontologies cannot in principle be aligned, just subsumed under the same core ontology, but brute reality in some domains (for instance traffic, taxation, consumer law, etc.) presumably can, certainly in very similar economies and cultures with shared values and concepts. This however creates the modeling problem of *grounding* institutional reality in a model of the legally relevant consensus reality (cf. [94, 63, 92, 93, 265]).

Section 6.3.1 however argued that specificity on a functional level is not the same as ontological subsumption if we combine ontological subsumption with ampliative reasoning. The law is organized around the behavioural *scripts* it attempts to influence: specificity in the context of normative conflicts, and “interesting similarity” in comparative law are not merely found in subsumptive relations between scripts. Some things do not subsume each other but are nevertheless considered sufficiently similar if a certain *abstraction* would be applied to them.

The comparison reveals differences and similarities relative to the evaluation dimension created by legal norms. Some similarities and differences are however more interesting than others: not all of them are worth reporting. Whether they are worth reporting depends on norms of analysis: an evaluative difference is interesting if the norm of analysis applies to it. If the norm of analysis postulates $|C| \prec_{na} |C'|$ while jurisdiction j_1 confirms it, $|C| \prec_{j_1} |C'|$, and j_2 is contrary to it, $|C| \succ_{j_2} |C'|$, the difference between j_1 and j_2 is interesting.

What is in effect compared in a rhetorical comparison is one script relative to two legal systems, with the same set of subjective and communal norms in the background. The fiscal examples above for instance relate mostly to investing: there is a difference between money in one’s hand – tax is to be paid in the form of money – and the estimated monetary value of capital investments in shares, bonds, real estate, etc. One can move money into investments, producing a capital yield, or move it out and keep it at hand. There are tactical considerations connected to the timing of moving it in or out, and it is for instance undesirable if one must sell a house at a certain point in time to pay taxes over its possession.

Regardless of whether the context is comparing alternative sets of rules during legislating, or comparing two different solutions in different jurisdictions, the differences are analyzed from the point of view of certain *stakeholders* and their (subjective and communal) norms (see section 4.5). These stakeholders are not persons but agent roles that participate in certain salient scripts. The normative systems are analyzed by looking at the choices they create in prototypical executions of a script, and whose interests are being served by those executions. The resulting arguments – *this agent role in this setting has a better alternative available relative to n_a in j_1 than in j_2* – are then in some cases generalized to populations – *up to x people may be affected negatively by adopting the rules of j_2* .

The question is where one gets these scripts. In the case of fiscal law these will be supplied by economists, together with their norms of analysis. In many other cases it is obvious to resort to case law.

Court decisions are phrased in terms of factors relevant to the system in which the

decision was taken. The court decision will not contain the information necessary to decide whether the factors relevant to the decision in the other system are present or absent. Using case law from legal system j_1 to explain differences with legal system j_2 results in a representation bias towards system j_1 . The differences that will be found can only be found relative to the contrasts that j_1 creates but j_2 doesn't.

Using case law from the old system to predict the performance of the new system is equally biased. One can only verify that it improves on the old system with respect to important contrasts relevant to the old system.

The same applies to a "migration of persons, goods, and services" scenario. In section 6.7.1 the example of web shopping was discussed. Many people may avoid purchases in other jurisdictions because of a perceived lack of legal protection. EU harmonization of the degree of protection only has a limited effect on people's attitudes, however, since what people note is first and foremost that the relevant scripts that they know about do not apply abroad. The difference in price hardly justifies having to learn something new.

In some cases the problem of comparing normative orders is contextualized to a clear setting that supplies the script. The growing interest in internet portals and web services is a good example: comparing for instance value added tax regimes for the purpose of creating a single calculation web service for different jurisdictions or an integrated portal for spatial planning restrictions from different jurisdictions (cf. [40, 45, 46, 41]) evokes clear scripts. The LKBS of course also embodies a set of scripts.

In this book separating the legal rules from the various assumptions about the quality of the available knowledge, the conceivable alternatives, the preferences of the user, etc., is a reusability issue for the legal rules. At the same time these other legally relevant knowledge sources also are also components of a reusable script when made explicit.

The conception of comparative argument as starting from a selected set of scripts is uncommon in legal knowledge engineering. With the benefit of hindsight the evaluative perspective of [42] *is* the script to be evaluated against. As noted in section 6.7.1, the script, as a typical mental concept, is in knowledge engineering represented by an abstract object whose structure is motivated by practical software engineering concerns, often without a genuine interest in the domain of cognition.

6.8. Conclusions

In chapter 5 I nearly completed my exposition on the representation of the contents of sources of law. The law consists of interrelated institutions which can be described as a system of legal rules and other legal facts. The sources of law should be considered a paper trail left by its operations, which often involve formal acts, for instance for legislating and for dissolving disputes in court.

In this chapter a few modest additions were made: section 6.2 introduced some more specifics on normative rules, and section 6.6.3 introduced a specific kind of applicability rule, the *choice rule*. The addition of these two rules finishes the construction made in this book of the problem-setting-independent and, in this sense, reusable knowledge of sources of law.

To explain the use of the choice rule I had to introduce *normative conflict*, which explains why choice rules are sometimes introduced for pairs of normative rules. The applicability of the choice rule is however not dependent on the presence of a normative conflict. Neither is normative conflict a phenomenon that arises from legal rules as such. Normative rules are defeasible by applicability rules and requirements posing additional constraints on the cases qualified by the normative rule; Since the normative rule only derives *its* violation, normative rules are not normally defeasible relative to other normative rules.

The issue of normative conflict lies beyond the institutional interpretation of legal rules; To explain it, one needs to appeal to a normative order intended by the legislator and show that the normative rules do not result in it for a specific category of cases.

Both the issue of normative conflict and another phenomenon of interest in legal knowledge representation, the contrary-to-duty norm, are best explained in terms of obligation or subjunctive betterness. Some *collections* of normative rules – think for instance of a single criminal code – are expected to result in a coherent account of normative order in terms of a *set* of propositions about a betterness relation partially ordering alternatives. If this account is inconsistent, it exposes normative conflicts. If the explanation of a normative order as a single set of equally important propositions fails for the analyzed collection, an alternative conceptualization of the collection must be found. In literature we for instance find the notion of a partial order on normative rules, but this account is not entirely satisfactory in cases of *intersection conflict*.

Normative conflicts are eventually resolved by explicitly formulated choice rules formulated for the occasion in which the normative conflict arises. The courts are the legitimate authors of choice rules. The mechanism by which a court arrives at a relevant choice rule for a normative conflict is not one that should be integrated into the OWL DL representation of a source of law, for the simple reason that the explanation of the normative conflict may appeal to making a closed world assumption about the set of alternatives available to an agent: the court only considers conceivable, i.e. known, alternatives.

Since the legislator is neither omnipotent nor omniscient in his abilities to bring about normative order in brute reality, his attempts to do so will inevitably fail occasionally. *Intentions* have a habit of coming to the foreground when attempts and failures matter: courts appeal to the intentions of the legislator to explain how legislation should be applied. Perceived failures of the legislator – commonly in the form of perceived normative conflict – create the space for courts to get involved in creating normative order, by formulating choice rules, which are therefore often found in case law.

Normative order obviously plays a broader role in problem solving than merely resolving normative conflict. Resolving normative conflict is not its primary function.

The rules of the institution have an important, instrumental function for those that engage in social interactions governed by those rules that goes beyond their interpretation as institutional qualifications of behaviour. Reciprocal recognition of actions by persons depends on the adoption and attribution of agent roles. The legal institution creates and refines such agent roles through its rules, *and* gives people reasons to adopt and to attribute them.

These agent roles play a central role in planning and plan recognition for the addressees of the law and the people that interact with them. All legal rules tend to have this effect, if they are known and understood.

Normative rules certainly give people a reason not to perform some acts, and a reason to expect others to do so, but the importance of normative rules as a guiding force in behaviour is perhaps overrated in computer science & law. The reason for this attention to normative rules is a simple and obvious one: the account of normative order in terms of obligation is relatively well-understood and obviously fits well on sets of normative rules, and normative rules are a characteristic feature of legal institutions. The normative order, as perceived by its participants, is however not an account of a system of normative rules in terms of obligation.

The Hohfeldian relations introduced in chapter 3 offer a richer vocabulary, by also distinguishing relevant relations based on the *power* to create certain legal facts, but it is only the fact that another agent takes an interest – exhibits preferences towards them that are not directly based in normativity – in these legal facts, that makes the agent recognize a legal relation.

I use an account based on a betterness relation to explain normative conflict – perhaps more accurately the *dilemma* some sets of normative rules pose in planning – within a set of normative rules presumed to express a coherent normative order, while I also suggest that the preferences attached to agent roles can be described both in terms of such a betterness relation and, if so desired, even in terms of normative rules. In this case the agent role takes the place of the norm or collection of norms.

Perhaps it seems strange to do so, but if the normative rules the legislator formalizes give rise to both a reason to make choose certain alternatives over others, and a reason to predict that others will do so, then the normative rules can be used for both a prescriptive and descriptive function.

As MacCormick pointed out (cf. [192] and example 6 in section 4.4) this is in fact the strange thing about many applications of the idea of institutions, beyond its application in law and rule-bound games: the participants in “institutions” often *cannot* in fact verbalize the rules in effect, which is a reason to doubt the very distinction between descriptive rules that merely describe regularities in behaviour and constitutive rules that bring about those regularities.

In the end these different accounts are accounts of goal-directed behaviour linked to *agent roles* assigned to a person.

The final section of this chapter returned on an application context of legal knowledge representation discussed only in the introduction and section 4.5.1.

The values the legislator tries to promote through the intended normative order unsurprisingly also take the form of propositions about a betterness relation, the *norms of analysis*. A weakness in the prevailing view of comparing legislation based on ontology integration is that it completely ignores the legislator’s intention to change an existing normative order. Comparative law in actuality compares intended and actual normative orders *in in the context of relevant social scripts*, and not only the part of the script explained by normative rules.

If knowledge engineers limit themselves to the institutional interpretation of legal rules, no meaningful comparison can be made by a KBS, because no knowledge is

available about which institutional entities and rules perform substantially the same functions in different legal institutions. The interpretation of normative rules in terms of obligation or betterness, without reference to certain specific social scripts that function as context of the comparison, will be of limited help, since, firstly, legislators use other instruments for influencing behaviour besides normative rules, and, secondly, not only the law is a reason to behave in certain ways. Moreover, intentions with respect to normative order are ascribed to the legislator: a better authoritative source than the source of law on the intentions of the legislator does not exist.

Chapter 7

Summary and Conclusions

7.1. Introduction

This book aims to be a work of ontology: an account of relevant aspects of the knowledge domain of law from the perspective of a legal knowledge engineer interested in sources of law. One cannot however say that the result of this work *is* an ontology; This book presents a mix of design principles, design patterns for knowledge representation in OWL DL, and ontology fragments.

Section 7.2 summarizes the essentials of the conceptual framework constructed throughout this book.

MetaLex and the Legal Knowledge Interchange Format (LKIF) ontology played an important background role. The notions of representation and reference (to concepts through terms) were introduced by me in MetaLex, and betterness in relation to normative rules was introduced by me in the LKIF ontology. Three concepts that play a central role in this book are *not* part of either ontology: the notion of *constitutiveness*, the notion of *applicability*, and the notion of *intentionality* as the *execution* of a plan. Also the notion of playing an *agent role* is not found there.

Section 7.3 focuses on a number of conclusions, both positive and negative, that can be drawn from the experience of constructing this conceptual framework. These conclusions address the questions of section 1.3. Section 7.4, which ends this exposition, addresses a number of interesting avenues for future research that are beyond the scope of this work as outlined in chapter 1.

7.2. Summary

Based on experiences over the last years with the MetaLex standardization effort and with the Legal Knowledge Interchange Format, this book discusses a number of relevant features of legal reasoning, and approaches for isolating different types of legal knowledge.

The legal knowledge source par excellence is the formal (written) *source of law*, discussed in chapter 5. The formal source of law is a bibliographic entity: a document. It exists in different ontological strata.

It is initially a physical *item*, that can be copied into additional items, which are however still examples of the same document *manifestation*. It can be edited, without changing the information represented by the document, producing additional manifestations of the same document *expression*. Even the information contents of the document can be changed, producing alternative expressions, while it still remains the same *work* from the perspective of bibliographic identification. The work level is only relevant if there is common content in all expressions of the work, and if later expressions share bibliographic identifying data with the original expression.

The XML standards on which the Semantic Web is built tend to identify documents on the item and manifestation level, while the legal field identifies documents on the expression (consolidated versions) and work level. For knowledge representation the expression level is the level of interest, as it is the level at which documents are identified and individuated by their information contents.

Chapter 4 described law in terms of institutions whose primary purpose is to create normative order by way of formalization – the sources of law specify a *formal, institutionalized normative order*. An *institution* is a *social* artifact: the structures of the institution are defined by the institutional facts that make up the institution, and its mechanisms of change are the constitutive rules that specify what *constitutes*, or counts as, an institutional fact.

The source of law is a writing that can be, is, was, or presumably will be used to back an argument concerning the existence of a certain institutional entity in a certain legal institution: The source of law is the result of a legislative act performed with the intent of creating that institutional entity, and functions as evidence of the legislative act.

The *legislative act* belongs to a broader category of *legal acts* that are characterized by 1) the requirement that one *intends* to bring about a certain institutional change, and 2) that this intent is communicated in writing. The legislative act distinguishes itself by the nature of the institutional entities it creates: legal rules. Other legally relevant acts also change legal institutional reality, but do not create recognized legal rules. Legislation may also achieve some other auxiliary effects, relevant to the operation of legal rules. The opinion that sources of law represent legal rules is so widely shared that it was uncritically adopted in this work.

The main function of the source of law is to function as evidence that certain institutional events, which changed the state the institution is in, happened. The sources of law function as a log book of relevant legislative changes to a legal institution.

A well-picked body of sources of law may also be considered a blueprint or a snapshot of the rules and structures of a specific legal institution of interest in a certain time frame, if the legislator does his work as a designer well. The state of the documentation and the state of the institution do not however necessarily simulate the other in either direction, as section 5.2 showed, even though the information contained in the documentation is essential for the social recognition of the institution. The sources of law and the institution, and the legal rule and the specific text fragment used as evidence for its existence, should be distinguished to avoid unnecessary complications in version management.

The task of the legal knowledge engineer, who is interested in representing legal knowledge, is primarily to reconstruct and represent the rules and structures of the institution rather than the source of law, which is itself a representation of (changes to) the rules and structures of the institution.

7.2.1. Reference and Representation

A specific goal of the recommendations in this book is liberating knowledge engineers from the self-imposed straightjacket of trying to represent a legal rule by a logical proposition. The logical propositions are statements *about* the legal rules and other relevant entities in the institution. A good argument against the bijective mapping of legal rules to logical propositions was found in [195] (in the context of norms; cf. section 2.2.2): legal rules have no truth value and truth-functional connectives should not be applied to them.

Representation of a source of law expression thus consists of two initial interpretative phases, although it doesn't end there:

1. Identifying the entities in institutional reality *first* represented by the source of law, as these were initiated by the legislative act evidenced by the source of law; and
2. formulating the logical propositions that constrain the valid models of institutional reality and describe the interfaces by which institutional reality can be changed, as evidenced by the sources of law and the way they are used.

The knowledge representation describes institutional entities as statives – entities whose existence is limited in time – linked to the sources as evidence of their existence through the mechanisms of *reference* and *representation*.

Section 5.4 proposed a number of principles for identification of institutional entities in sources of law. The source of law at the expression level is identified by a URI, *refers* to a set of URI-identified *terms*, and it *represents* a set of URI-identified *legal rules* and other *legal assertions*. This applies to the source of law as a whole, and to any URI-identifiable fragment of the expression. Since the expression does not change – it is only created – one would expect that the set of legal rules it represents and the set of terms it refers to also normally do not change. In the rare case that they do, we are dealing with an *ex tunc* expression, which is the result of a legal fiction.

Terms can be categorized into those that are *in* the institutional reality created or modified by the legal rules, and those that are relevant to this institutional reality but outside the boundaries of the institution. Legally relevant terms are not defined by the institution: whether the term does or does not accurately describe a certain entity is from the point of view of the institution a *technical question* (cf. section 2.5).

All legal terms and rules *first* identified in an expression-level source of law should be assigned the same namespace, which is bijectively related to that source of law. No other terms are in that namespace. A new URI, in a new namespace, can be assigned to a rule or term only if a rule or concept is deemed to change in a new expression of a work, in the sense that the old concept and the new concept denoted by the same term in the source cannot be reconciled. The decision to integrate terminology from different expressions is made explicit in the form of equivalence statements.

One might think of the set of terms and legal rules occurring in all known expressions of a work as the shared set of terms and rules at the work level, but the shared work level set only exists from a specific vantage point in time, or only once the source of law has become immutable after its repeal, and the shared work level set – which can no longer change – has become largely irrelevant. This abstraction cannot be safely used, since it is likely to lead to versioning problems.

The source of law on the expression level however *cites* other rules on the work level, while the legal rules we represent are necessarily identified by their *first representation* in an expression, and not on the work level. A citation (*text fragment*) *w* *applies to* (*concept*) *C* should for instance be read as *each legal rule that is represented by an expression-level text fragment that embodies work fragment w applies to C*.

7.2.2. Legal Rules and Logical Propositions

Relevant patterns in logical propositions describing the functions of legal rules revolve around the notions of *constitutiveness* and *applicability*. In particular the notion of constitutiveness, introduced in section 4.2, is central to the institutional interpretation of law.

The legal rules represented by the source of law appeal to two separate realities – institutional reality and brute reality – and perform a mapping from brute reality – the ontological substratum – into institutional reality – the ontological superstratum. The substratum has an existence independent of the rules, while the superstratum is supervenient on the substratum and exists by virtue of social recognition of the rules of the institution. Institutional events are *constituted by* events in brute reality.

Institutional rules map out a logical space of possible models of the institution: they form the institution’s ontology, and can be interpreted as terminological axioms. They are not defeasible as a matter of convention: if the suspicion arises that a certain institution has an ambiguous internal structure, this should serve as a hint to either redefine the boundary between the institution and brute reality, or to split the institution into two or more institutions.

Since the quality of the mapping between brute and institutional realities is likely to be imperfect, we should assume – as a matter of convention – that institutional facts only exist as long as it is consistent to believe they exist.

The main function of the *constitutive rule* is to define the interface through which the state of the institution can be changed. As such, the most important question they answer is how to map brute reality *into* institutional reality. Legal rules of a constitutive nature are however not conditionals which take brute reality propositions as a premise, and institutional propositions as conclusion. The constitutive legal rule may have a variety of functions, corresponding to certain patterns of logical propositions describing them.

Logical propositions *describing* constitutive legal rules (cf. sections 4.4.1 and 5.2) can be categorized into *requirements*, which posit necessary conditions on the mapping between brute and institutional fact, and *indicators*, which license an inference from brute fact to institutional fact, if and as long as it is consistent to do so. In addition, the law may impose constraints on knowledge availability: there must be positive justification from brute reality – not from the legal rules alone – that a change to institutional reality happened. One cannot freely infer institutional facts from lack of knowledge. These rules were tentatively called *burden of proof* rules. In an argumentation context, these function as (epistemic) obligations to justify arguments with certain information that must be collected (usually by questioning a user).

Requirements, and the terminological axioms used to describe institutional rules, can be accurately captured by straightforward OWL axioms. Indicators and burden of proof rules are however autoepistemic in character.

Applicability, mainly discussed in section 5.2.2 and 5.2.3, plays a central role as soon as the logical proposition and the legal rule are separated. If a logical proposition states that a certain proposition indicates or entails another proposition, we must add to this conclusion that a legal rule has been applied. Any logical proposition about

the application of a legal rule must conclude that the legal rule has been applied.

A special class of legal rules (generally described as requirements) constrains the applicability of other rules. *Applicability rules* are used to avoid repetition of the same requirement, to demarcate the extent of jurisdiction over people, territory, and substance, and applicability in time, and to restrict the applicability of legal rules made by a lower legislator. The *choice rule* makes the application of one legal rule conditional on the application of another legal rule, if the combined application of two legal rules is judged to be logically ambiguous or *contrary to the intended normative order* (cf. section 6.6).

Applicability rules can be used as a guide for clustering sources of law into coherent domains of application, and are generally used for that purpose by the law. They should also play that role in legal knowledge engineering, but generally do not do so explicitly. This book explains legal fiction, and the closely related also-applicability pattern, in terms of applicability: legal fictions are inconsistent with the conscientious use of applicability rules (cf. sections 5.2.4 and 5.4.3).

Sections 4.6 and 6.2 also distinguish the *normative rule*: it is a constitutive rule characterized by the type of institutional fact it establishes and the intention with which it is created. Its logical pattern does not *have* to be distinguished from the already established forms if one doesn't attempt to represent the *intended normative order*.

This observation sets this book apart from previous uses of Searle's constitutive rule (cf. [249, 248]) in legal knowledge engineering. Although Searle subscribes to the logical distinction between constitutive rules and "regulatory" or normative rules, the *legal* normative rule should clearly be considered a constitutive rule (cf. section 4.6). Treating only *some* legal rules as constitutive only confuses the issue.

Note that the logical propositions following the identified patterns are not found in the source of law: they are created by the knowledge engineer. The logic of grouping URI identifying logical propositions into namespaces, and the sentences themselves into files, and doing version management on those files, is a knowledge engineering concern not constrained by the design rationale of the sources of law. The representation proposed in chapter 5 can be expected to be compatible with other methodological concerns besides isomorphism with the source.

7.2.3. Normative Order

The work of the legal knowledge engineer is not limited to reconstructing and representing the rules and structures of the institution. One could argue that the task of creating a reusable knowledge component out of a source of law ends there, but the resulting knowledge components have little purpose if not augmented with another interpretation at yet another level of existence: the *intended normative order* (cf. chapter 6).

An institution can intend and attempt to formalize a normative order, but the existence of normative order itself is only proven by regularities in the behaviour of individuals in a community.

Normative order always exists, regardless of whether it is formalized, and the

legislator usually just tinkers with the details of an existing normative order in society. Sometimes the actions of the legislator only confirm, or perform minor repairs in existing, spontaneously arising normative order, and sometimes the legislator only tries to create the right conditions so that acting in one's own interest will create the intended effects.

Even for the mere purpose of complying with the law, we have to ascribe to the legislator the intent to formalize a specific normative order. The *normative rule*, discussed in section 6.2, has a straightforward translation: what it *tells* us is that a certain behaviour brings about the institutional fact of violation of the normative rule, and what it *means* in terms of the intended normative order is that the legislator intends us to choose against execution of plans that could be recognized by others as a violation of the normative rule: one could call this the *deontic choice* effect of the rule.

Observe that deontic logics represent not the legal rules, but the intended normative order as readily apparent from the normative rules. As observed in sections 4.6 and 6.2 a straightforward assessment framework classifying cases as allowed or disallowed suffices for the institutional meaning of the normative rules of the institution. Since the sources of law do not formalize the whole intended normative order, a representation of these rules in deontic logic gives us only a partial account of behavioural regularities in a community.

The legislator acts on an existing normative order, and performs legislative actions with the intent to change it. The legislator relies on three important mechanisms to effect changes in the normative order: public announcement of the rules and structures of the legal institution and the competence of addressees of his announcements to understand the intended implications for the normative order, the widely held preference for predictable social interactions among the addressees, and the ability of the widely recognized legal institution to effectively organize punishment of those that don't comply with the prevailing normative order, which underlies the deontic choice effect of the normative rule.

That the ascribed intended normative order – the whole of it – has direct relevance for the application of normative rules becomes clear in situations of normative conflict and compliance dilemma: if compliance with the normative rules fails to achieve the intended normative order, this calls into question their *applicability* to the case even if there is no plausible argument that application of the rule would create an inconsistent institutional reality (cf. section 6.2, specifically 6.3.1).

Just like the sources of law cannot be taken at face value as a blueprint of a legal institution, the normative reading of its rules cannot be taken at face value as a blueprint of the normative order it formalizes.

This is the most important reason why in the field of comparative law the importance of taking into account the cultural and social context in which legislation exists is often stressed, and why *legal transplantation* of legislation doesn't always have the expected effects (cf. section 6.7.3). Filling in the gaps is an interpretative task, and involves ascribing a clear intent to collective agents – parliaments, governments – that very likely had no such monolithic intent in the first place.

In many cases presuppositions about the pre-existing normative order being re-

gulated are found directly in the source of law: Hohfeldian transactional notions like rights correlative to a duty, liability correlative to a power, etc., are founded in certain expectations associated with the adoption of certain agent roles in existing social scripts (cf. sections 4.8.2 and 6.7.1). The legislator may for instance choose to phrase an obligation in terms of a right of a counterparty; This does not change the fact that the rule should be interpreted primarily as an indirect obligation, but it does yield information about the intended normative order. Another clear source of presuppositions about the intended normative order is found in the evaluative nature of the terms used: legal qualifications are often either value-positive or value-negative (cf. section 6.7.1).

Legislative explanatory documents accompanying legislation also often explain the intended normative order. Articulations of the intended normative order play a direct role in the form of *norms of analysis* in the legislative drafting phase (cf. section 6.7.3). As argued in section 6.7.2, courts also ascribe an intended normative order, and latch on to deficiencies in the operation of the rules relative to this ascribed normative order to create legal rules of their own: applicability rules which make the application of one rule conditional on the application of another rule if certain conditions are met.

Consider also the perceived problem of giving John permission to live in a certain house and Jane permission to destroy it, discussed in section 6.3. It is a problem only because we immediately ascribe to John a preference for i.a. staying alive, and to the legislator awareness of John's preference and the willingness of taking into account John's preference for staying alive: the legislator is therefore behaving irrationally in our opinion in his attempts to create normative order.

As observed in section 6.7.3, analysis of comparative arguments about differences and similarities between legal regimes has the potential to teach us the difference between the institution and the normative order it proposes, and requires a great degree of intercoder reliability in the representation of normative order: legal knowledge engineers can use this information for more robust legal knowledge representation. At the same we have to recognize that there is little demand for LKBS that evaluate proposed legislation, and little prospect for investment in the development of such LKBS. Comparison for the purpose of assisting the mobility of people, goods, and services in the EU, or for preparing the implementation of legislative change in administrative settings *is* on the other hand of recognized relevance to society.

7.2.4. Subjunctive Betterness

In the representation of normative order the *subjunctive betterness* conditional plays a central role in this book (cf. section 4.6 and chapter 6 generally, section 6.2 specifically). In OWL, the legal normative rule of obligatory form $O_n(\alpha \mid \beta)$, where n is the reified normative legal rule, can be represented as follows:

$$\begin{aligned}\alpha \sqcap \beta &\sqsubseteq \forall \preceq_n \alpha \sqcup \neg\beta \\ \neg\alpha \sqcap \beta &\sqsubseteq \exists \preceq_n \alpha \sqcap \beta\end{aligned}$$

If the obligation is for instance *when driving a vehicle one should keep to the right of the road*, the interpretation in terms of subjunctive betterness is:

When driving a vehicle and keeping to the right of the road, all equal or better alternatives would involve not driving a vehicle or keeping to the right of the road. When driving a vehicle and not keeping to the right of the road, an equal or better alternative would be driving a vehicle and keeping to the right of the road.

The combination of both statements is only satisfied by considering alternatives involving not keeping to the right while driving strictly worse than those that do involve keeping to the right, or by not considering alternatives which involve driving a vehicle at all.

Preferences – whether communal or private – can be represented in the same way (cf. section 6.7.1). Normative order arises from the fact that a large body of such statements – representing legal, communal, and private normative rules – are generally accurate descriptions of the behaviour of a community because people tend to copy both the behaviour and the justification of that behaviour of those around them.

This book also sketches how this information about the prevailing normative order is organized and applied (cf. section 4.5): Kahnemann’s reference transaction is a script for getting something done using other people as a resource. The script generally works because the other participants act in accordance with their agent role in the script: they understand and subscribe to the prevailing normative order. The *agent role* functions as a anchor for ascribed preferences: someone is not considered a *seller of snow shovels* because one merely has the competence or ability to do so, but because others expect that they can engage with this seller in a predictable transaction for obtaining a snow shovel in exchange for some predictable amount of money.

Unfairness is deviation from the scripted execution of a transaction to the advantage of another participant and to the disadvantage of you, without an acceptable justification.

As long as we accept that a large body of such subjunctive betterness propositions are shared, and that propositions like *I prefer that you . . .* are irrelevant to describing a prevailing normative order, subjunctive betterness propositions generally do not have to be indexed to their origin (i.e. we do not have to distinguish between those in your and those in my mind). Even the ones of legal origin do not have to be distinguished as to their origin, as long as their context of applicability is suitably constrained to time period and jurisdiction by applicability rules.

7.3. Conclusions

For the purposes of maintenance of legal knowledge-based systems, a one-to-one mapping between knowledge components and sources of law, and between fragments of text in the source of law and logical propositions, would superficially appear to be ideal. This expectation is based on the assumption that the legislator’s design rationale for sources of law and the knowledge engineer’s design rationale for knowledge components can be functionally aligned. The Semantic Web vision moreover suggests that such knowledge components are reusable, or can be designed for reusability.

The problem of making high-quality reusable knowledge components representing sources of law available on the Semantic Web can be approached from several angles, as indicated in the introduction of this book.

7.3.1. The Semantic Web

In section 1.3 a technical theme was addressed: how does one integrate existing core technologies and technical standards of the Semantic Web, like its naming and addressing standards, and the description logic OWL DL, with legal knowledge representation?

Of central importance is the uniform resource identifier or URI, and the realization that the URI violates the unique name assumption typical in knowledge representation. This has consequences for inference in OWL DL, as explained in section 3.4.

OWL DL is a monotonic description logic, designed specifically for robustness in a distributed and fault-tolerant environment. Description logic is particularly useful for describing terminological axioms that are true by definition. It however lacks the possibility of expressing propositions about propositions, and does not support non-classical, but for KBS development important, features like defaults, integrity constraints, and negation as failure. These are certainly necessary for intelligent behaviour, as section 2.4.1 explained. Arbitrary use of such features would however certainly not help to achieve our goal of minimizing impact in the deployment of updated knowledge components.

Chapter 3 sketched a general layout of straightforward knowledge-based systems using OWL knowledge components, that accounts for the controlled use of such constructs. A distinction was introduced between a number of types of logical proposition: OWL DL's terminological axioms and assertions, on the one hand, and default rules and integrity constraints on the other hand. This distinction describes several existing methods for interfacing KBS to OWL DL. Moreover, many proposals to extend OWL DL focus on such constructs.

The very general level of abstraction at which these constructs were described, permitted description of the relation of such constructs to particular types of legal rule without undue epistemological commitments to a specific problem solving strategy.

7.3.2. Ontological Stratification

The great challenge is to structure legal knowledge in accordance with the distinction made between types of logical propositions, and into coherent, justifiable, and reusable, knowledge components. Chapter 2 introduced some theoretical devices for this purpose that classify knowledge by function. To give a philosophical justification for localizing the use of default rules and similar defeasible constructs, a number of concepts were introduced: abstraction, aggregation and refinement of abstractions, and ontological stratification (cf. section 2.3.1).

To effectively apply these theoretical devices to the domain of law, and specifically sources of law, the inspiration must be found in legal theory (generally chapter 4).

Knowledge in its traditional sense, as attainment of truth, usually describes “islands” of abstract knowledge that can be considered axiomatic: *if* an abstract theory is applicable, it is indefeasibly applicable as a whole, but *whether it is advisable to try* depends on the thing one want to apply the theory to, and on the problem one is solving. Abstract theories always come with if's and but's. Terminological axi-

oms accurately model these theories, while default rules function as indicator that the abstract theory *may* be applicable.

This philosophical justification accounts for defeasible rules, for problem solving techniques like heuristic classification, etc. These are alternative epistemological accounts of the same observations about the character of reasoning with defeasible knowledge.

The nasty details are in the abstraction moves between abstract theories: an abstraction is only permissible as long as it remains consistent with the current belief set, and only useful depending on the problem one is solving. It is these abstract theories whose terminological core can be expected to remain stable over long time periods. These “islands” of stable terminological knowledge, surrounded by rules recommending a jump to the island, are what is captured by a reusable knowledge component.

The concept of an ontological stratum suggests one should look for ontological criteria to tell apart the abstraction and the thing it is an abstraction of. This book proposes several abstractions between ontological strata relevant to law, based on an analysis of central concepts in legal theory relevant to legal knowledge engineers, notably constitutiveness, execution, and normativity (generally chapters 4 and 6), and proposes that defeasibility is limited to such abstractions.

7.3.3. Reusable Knowledge Components

When designing standardized components it obviously makes sense to focus on the functions of the component; These characterize how one interacts with the component. These characterizations of function can be identified at different levels of abstraction, as indicated, and depend on the kind of purpose one has in mind. This of course also applies to knowledge components, and to the source of law.

In this book we have focused on different levels at which we can characterize the source of law (generally chapters 4 and 6):

1. as an account, physical evidence, of legislative changes to legal institutions, and
2. therefore as an indirect and incomplete representation of the structures and the interface of legal institutions, or
3. as the physical product of acts intended to change prevailing normative order towards an intended normative order, and
4. therefore as an indirect, incomplete, and approximate representation of normative order.

This book concludes that constitutiveness and normativity are *different accounts* of the functions of the source of law, *not* alternative functions of legal rules as Searle’s distinction between constitutive and regulative rules, taken up in legal knowledge representation literature, appears to suggest.

The obvious choice is to characterize the function of the source of law, and therefore of the knowledge component, in terms of the legal institution(s) it provides knowledge of. This for instance means that even the normative rule must be in first instance accounted for as a constitutive rule.

The relation between the source of law and the institution is clearly a more direct one than the relation between the source of law and normative order; Effects on

normative order are rather an effect of the legal institution and its social recognition than of the source of law *per se*.

A second reason to discount the relation between source of law and normative order is that law only changes a pervasive, existing normative order, and not always in predictable or well-understood ways.

A third reason to discount the strength of this relation is that even the *intended* normative order, the thing the legislator's actions are aimed at, is at best something to be found "between the lines" in the source of law, if the ascription of such an intended coherent normative order is an accurate model of legislating at all.

The notion that the sources of law give a full account of the structures and rules of the institution is only tenable to the extent that the legislator provides full evidence of it. Since the legislator does have good reason to provide such an account, this concept will work reasonably well. There are however occasional counter-examples, as indicated in chapter 5.

7.3.4. Isomorphism with Sources of Law

The relation between the source of law and the knowledge component is of course easier to establish than the relation between a knowledge component and an institution. The source of law may account for multiple institutions, and generally does so only partially.

The traceable relation between the source of law and the knowledge representation is not found in a correspondence between logical propositions and fragments of the text. This position is suspect philosophically *if* it means that these logical propositions must be connected by truth-functional connectives (which is by the way not the only account one could give of them), and – more importantly – it is highly impractical in a language like OWL DL that limits propositions about propositions.

The position taken in this book is that the source of law *refers* to concepts (through the use of terms) and *represents* legal things, amongst others legal rules. Legislative acts create, change, or destroy these things, and in addition produce physical evidence and representation, in the form of the text of the source of law, of these creations, changes, and destructions. The text is constitutive of the legal thing.

Adaptation of the knowledge representation to changes of the underlying sources of law relates to change of these (institutional) concepts and the legal rules, and not to logical propositions.

Both terms (or concept names), and individual objects are uniquely (but not exclusively) identified by a URI in OWL DL axioms. Although logical propositions about these named entities *can be* identified by URI in OWL DL, this URI is not accessible for reasoning within the *same* OWL DL theory. Legal rules will however directly refer to text fragments in sources of law, in order to refer to other legal rules. Section 5.4.1 gave an example of how this is addressed within the same OWL DL theory.

Traceability to logical propositions is provided by the principle that the logical proposition is *about* the application of a legal rule to some thing (through the **applicable** property).

7.3.5. Legal and Legally Relevant

The analysis of sources of law in terms of institutions supplies us with a criterium to tell apart legal knowledge from legally relevant knowledge, which was one of the objectives set for this work in section 1.3: legal knowledge concerns knowledge of institutional concepts, institutional facts, and institutional rules, and the constitutive rules that tell us how brute facts are abstracted into institutional ones.

Legally relevant are the concepts that describe the brute reality that is abstracted into legal institutional reality by constitutive rules, and – arguably – various epistemological theories that help explain how law enters the picture in planning and plan recognition, how planning dilemma’s should be resolved etc. Legally relevant knowledge is not found in the sources of law, and therefore has no place in the knowledge components of interest in this book.

Constitutive rules do not only indicate, but also restrict, the possibility of abstraction into institutional reality. This is why a distinction between indicators and requirements was introduced. Constitutiveness is based on an ontological criterium: if the antecedent and the consequent of the rule are in different strata, the rule is constitutive.

This expands the notion of constitutive rule beyond the notion of *supervenience conditionals*, found in literature, which only explains indicators.

It has been claimed in legal knowledge engineering (cf. for instance [130], and the references given there in the appropriate section) that one of the strange features of constitutive legal rules and legal facts is that a legal fact can only come into existence if one of the constitutive legal rules that has it as a conclusion is applicable; This would appear to mean that, as eloquently formulated in [130]:

[..] two or more supervenience conditionals cannot all be seen as definitions at the same time. They rather seem to imply, together, a definition having as antecedent the disjunction of the antecedents of the different conditionals. If consenting supervenes on signing a document, and it supervenes also on signing via internet an e-form but on nothing else, then signing a document or e-signing imply consenting and conversely consenting implies signing or e-signing.

This clearly only explains something about indicators. Moreover the notion of a disjunction of the antecedents of the various conditionals is inherently problematic for maintenance and should be avoided.

This interesting feature of constitutive rules has not been reconstructed in this book through the use of logical form, but of an ontological commitment: each institutional thing is constituted by a constituting thing. In the case of the law, each legal thing is constituted by a legally relevant thing, to which a legal rule has been applied. That this mapping from constituting thing to institutional thing is only constructed through the use of the rules of the institution is a matter of social recognition: the rules of the institution specify its recognized interface.

[130]’s observation applies equally to for instance the classification of diseases; Only certain symptoms are recognized as indicators of a disease. In this respect the interface of the institution and for instance the interface of a component or for instance the user interface of a computer application are also similar: the interface specification is an

exhaustive description of the recognized – i.e. intended and expected – interactions you can have with it. This is also the case with legal institutions and their public interface.

Because the institution however only exists *because* of social recognition alternative interactions would only arise through social recognition of those interactions, while for instance a computer can be directly interacted with in unintended and unexpected ways with for instance a hammer.

7.3.6. Loose Ends

There are several loose ends in this proposal that may cause some uneasiness. Most important of these are the notion of boundaries between legal institutions, and the positioning of normative order.

The assumption that institutional reality *is* internally coherent, as opposed to the defeasible rule interfacing institutional reality and external reality, creates a new problem of identifying boundaries between different legal institutional realities.

External institutional realities that are recognized as legally relevant by a legal institution complicate matters, as these external institutions compete for normative order, may use the same vocabulary, and generally confuse matters for people that recognize the rules of multiple legal institutions and have trouble telling them apart.

The proposals in this book replace the problem of keeping track of defeasibility between potentially all legal rules with the rather arbitrary notion of splitting the law up into as many different legal institutions as required for delegating defeasibility to the interfaces between them. We for instance let go of the idea that one legislator manages one institutional reality: a legislator may unintentionally split one institutional reality into two by mere use of legislative technique.

The problem of telling apart institutions does not however appear to be *more* complicated than determining whether a rule may be defeated by another one. It supplies a new minimality criterium for defeasible reasoning in law.

Lastly, if legal knowledge is in essence limited to knowledge of the institution, then normativity is a mere derivative product of it. The rules of the institution give rise to preferences and to expectations about the behaviour of other people. This is not a special observation about legal knowledge: all knowledge has the purpose of giving rise to preferences and expectations, and all knowledge about social conventions and institutions give rise to preferences and to expectations about the behaviour of other people. There is no a priori reason why there should be a specifically legal account of this.

In law the applicability of legal rules may however depend on normative conflict, which can only be explained in terms of an account of normative order. Also the comparison between *similar* legal institutions, for instance from different countries is an ill-defined question if it does not include the comparison of two intended and prevailing normative orders. There are multiple ways to create the same normative order through different institutional rules and structures, if we take into account that the institution only tries to *change* an existing, prevailing normative order.

In chapter 6 a not entirely successful attempt was made to reconstruct the ma-

major categories of normative conflict in literature based on a simple representation of preferences in OWL DL. An abstraction to preference, or for instance to obligation, attributed to agent roles is justifiable for normative rules if we assume they are constitutive.

Agent roles play an important role in organizing out knowledge about behaviour, as pointed out in chapter 4: when one assumes an agent role, one assumes the preferences associated to it, and when one attributes an agent role to another one has expectations based on these exact same preferences as pointed out in chapter 6. It makes sense to assume that the preferences associated to an agent role should form a coherent whole, but a person may of course have to deal with the expectations based on different attributed agent roles.

Both the comparison of legal institutions and the resolution of normative conflict is hard to conceptualize without the context of specific scripts and agent roles, as chapters 4 and 6 pointed out.

While a construction of the intended normative order, in the context of the relevant scripts, is necessary to realize the wonderful promises of section 1.2, this construction cannot however be considered to be part of the knowledge components representing sources of law. This for instance also means that Hohfeldian concepts like right, liability, etc. cannot be completely accounted for; Whether one “has a right” depends on preferences and expectations, and not on a specific phrasing of a legal rule in a source of law.

This leads to an important conclusion that affects reusability concerns: there is a fundamental friction between reusability of knowledge components representing sources of law and their relevance; Knowledge about prevailing and intended normative order is more relevant than knowledge of the structure and interface of legal institutions, but

1. the sources of law do not directly contain it,
2. it is less reusable in the sense of being very contextual, and
3. prevailing normative order changes due to changes in society that are not mirrored by changes in the sources of law.

7.4. Research Questions

The law intends to formalize a normative order, but normative order only exists if a large community of people generally make decisions in compliance with the normative order. Complying with the intended normative order, or recognizing non-compliance is however not trivial.

Central to this problem is often the problem of recognizing intentions (cf. section 4.1). Every action is intentional, directed towards causing something, but there is a distinction between intentional and unintentional consequences of one’s actions. The concept of intention was explained in terms of action: we recognize behaviour as action when we explain it as a *plan execution*. Since we recognize that cognitive resources are limited, we assume that people remain committed to plans over a longer time period. The decision to execute a plan takes place in a certain discrete time period

and consumes scarce cognitive resources. Every change or non-change that was part of the plan, regardless of whether one is the physical cause of the change, and regardless of whether the change was part of one's objective in planning, is in a sense intended.

I made a distinction between behaviour descriptions in law that do (ought-to-do) and do not (ought-to-be) explicitly address an agent role and action (cf. section 4.7). Even if no agent role is explicitly addressed, this does not change the fact that only action violates normative rules: the subject of a norm or a preference is in essence always a situated action under conscious control, even if the legislator had no specific script in mind or didn't bother to mention it explicitly, as in many technical rules for design and construction activities.

In order to come to the conclusion that someone made the wrong decision, the outside observer will if necessary decide 1) what plan was being executed, 2) when the decision to execute that plan was made, 3) which alternative plans were considered, and 4) generally, what the decision maker knows in a dispositional sense. This *folk psychology* model of decision making presents legal knowledge engineers with interesting, and important, research questions:

1. I have pointed out a fundamental *deadline* problem in relation to many obligations: it is not possible to determine whether an obligation to bring something about has been violated without setting a deadline, if one separates the obligation from an ascribed planning activity (cf. section 4.7.2). In this case and in the case of *negative causation*, also briefly mentioned, one violates a normative rule because of the alternative plan one *did* execute instead of those that comply with the obligation: it is not the case that one violates a rule by non-action. If this were not the case, setting deadlines would receive a lot more attention in the field of law than it does. Negative causation, despite its name, has very little to do with causation. Both passive and active behaviour can be considered to be action, if it is interpreted as the execution of a plan.
2. In relation to compliance dilemmas we also encountered the question of whether there is an alternative behaviour that is allowed. As pointed out in sections 6.3.1 and 6.7, it is not sufficient to believe that there must be such an alternative; One must also *know* one to be confident in the decision that some behaviour is disallowed. The alternative we choose is the best one we know, not the best one there is. This applies not only from an internal perspective, but also to the external observer who ascribes a choice between concrete alternatives to the decision maker. The compliance conflict in literature would be easy to formalize in terms of a discrete and finite set of alternatives in a menu. When disconnected from a theory of decision making that supplies such menus, it becomes a very hard knowledge representation problem.

The meaning of normative order is inseparable from a model of planning. While knowledge representation is usually interested in intelligent planning activity, in keeping with the philosophy of representing *knowledge as a productive factor*, the law is concerned with *all* planning activity, not only particularly intelligent planning. AI planning literature has little to offer: the norm may play a variety of roles in planning models. Depending on the purpose of planning, the norm may be an ordering on al-

ternatives, a reason to reject alternatives, but also very often the apparent reason why alternatives are not even imagined at all. The behaviour of others is usually predicted in terms of the prevailing normative order: non-compliance is not even foreseen, and certainly not planned for.

There has been some initial work on the task of *legal planning* (cf. for instance [265, 27]). These however also conceive of legal norms as mainly as constraints in planning. The role of norms as enablers of action is rarely considered.

Connection of planning and situation recognition to Schank's scripts and Kahnemann's reference transaction (cf. section 4.5) as a device for explaining a prevailing normative order is still an open research question, and an interesting one. The problem of planning and ascribing plans to others should certainly be a central problem in legal knowledge engineering.

The notion of normative order also presents us with other interesting research questions of a more practical nature. It for instance makes sense that certain organizational notions like *business process* can also be considered as conceptualizations of a prevailing or an intended normative order within an organization (depending on whether one is modeling an existing one or re-engineering one). A mapping between the language of law and normative order and the language of business process modeling would have practical commercial applications in organizational settings for the coordination of organizational change and the revision of deployed legal knowledge-based systems in reaction to changes in the sources of law.

7.4.1. Abstraction

The mechanisms of abstraction, aggregation, refinement, and ontological stratification discussed in section 2.3.1 also present us with interesting research problems, but explaining these forms of reasoning is not specifically relevant to legal reasoning.

In this book, these mechanisms have at some points been used as an explanation, but how they arise and how they work remains unexplained. Seemingly arbitrary choices of granularity level play a prominent role in two common sense reasoning subproblems in legal reasoning. These both involve relating legally relevant terms through common sense knowledge not supplied by the law.

The first one concerns the question of whether the behaviour or state of affairs sketched in one rule is more specific than the one in another. This issue is of interest to the defeasibility of indicative rules in general, and to the resolution of normative conflicts specifically. Specificity appears to be much broader than subsumption, and it is a problem for nonmonotonic reasoning in general (cf. section 6.3.1). Descriptions at completely different levels of *abstraction* are sometimes said to be in a specificity relation to each other: the solution to this question is perhaps found in the equally mysterious notions of aggregation and refinement as applied to abstractions, in the sense that refinements, which are more often than not *imperfect reductions*, are more specific than the thing they are a refinement of.

Trying to emulate this form of reasoning is a considerably less efficient investment of one's modeling effort than representing the results of validated arguments of this kind in court decisions and administrative guidelines as choice rules.

A second, similar issue exists for choice of granularity in the description of actions and events. As noted in section 3.5.3, whether something goes through a series of changes or remains the same over a time period, is mainly a matter of selective attention. One may choose different ontological strata, different abstractions at varying levels of granularity, when describing a behaviour, but grounding every description encountered in a fundamental causal model of reality is not a task the legal knowledge engineer can set himself when representing a source of law.

Causal reconstruction is however of interest in certain areas of law, mainly criminal law and tort law, and support for creating causal models may find a niche market in this context. There are interesting efforts in this direction in [152, 184], and indirectly in [23]. In most applications of law, the recognition of actions is not considered problematic and the question of causality does not arise. This is because we already recognized behaviour as action, i.e. ascribed a plan, and implicitly attributed responsibility for certain events – and non-events – to the planner.

Both issues concern “brute” *legally relevant* terms, and the more general problem of grounding these in a fully developed model of the relevant parts of consensus reality, sometimes included in the quest for *deep models* or *deep structure* of law as for instance found in [94, 20]. This notion of deep structure, and the benefits of uncovering it, has been implicitly adopted in this work.

The notion of the intended normative order in this book is a deep model to be uncovered. Some deep models are however maybe better left unaddressed. The problem of providing a “deep” explanation of specificity and causality is in the opinion of this author too hard and in practice too insignificant to consider them as requirements for practical legal reasoning. The explanation of the deep structure of law in [94] (also adopted by, amongst others, myself in [286]) does appear to assume that a fully developed model of the relevant parts of consensus reality, including the specificity question at least, is a feasible project, but our work in the last decade has not resulted in a satisfactory approach to uncovering it.

Reconstructing the limited set of reference transactions the legislator had in mind when he chose a certain legislative solution *must* be feasible. Specificity is only interesting if we find it in the form of an implied subsumption in the language chosen by the legislator, which only happens if the legislator *designed* the choice into his legislation. In this case it is naturally addressed by a choice rule (cf. section 6.6). Construction of fundamentally new conflicts should be left to the courts.

Appendix A

About MetaLex and the LKIF Ontology

Many ideas put forward in this book have been presented in some form in other venues, although the material is mainly new.

Some ideas are based on my activities as chairman of the technical committee of the MetaLex CEN/ISSS workshop and in the years before while working on a proposed XML standard for sources of law (cf. [37, 38, 36, 51, 52, 49, 285, 48, 53, 54, 32, 282, 283, 190, 219]). Many of these ideas, mainly found in chapter 5, have been confirmed as a CEN/ISSS¹ publicly available specification (CWA15710).

Some knowledge representation ideas have been introduced by me in the *Legal Knowledge Interchange Format* ontology (cf. [35, 66, 153, 155, 31, 43, 44, 61]) of the Estrella project.

Both MetaLex and LKIF are publicly available as an OWL ontology: I have chosen not to include these as appendices, as they are works in progress and this book is voluminous enough as it is.

For more on MetaLex: <http://www.metalex.eu>.

For more on LKIF: <http://www.estrellaproject.org>

This book aims to be a work of ontology: an account of relevant aspects of the knowledge domain of law from the perspective of a legal knowledge engineer interested in sources of law. One cannot however say that the result of this work *is* an ontology; MetaLex and the LKIF ontology play an important background role in this book, since I introduced a number of concepts important to me in these ontologies, but the reader needs to keep in mind that both these ontologies were collaborative endeavors.

In particular MetaLex – a standard for legislative documents and metadata that has actually made it to the status of a CEN/ISSS prenorm, and is used by some computer applications – is very much the result of the work of a committee and as such represents only common ground between me and the other members of the CEN MetaLex technical committee. LKIF on the other hand sometimes appears to accommodate *alternative* solutions for solving the same representation problems.

Three concepts that play a central role in this book are *not* part of either ontology: the notion of *constitutiveness*, the notion of *applicability*, and the notion of *intentionality* as the *execution* of a plan. These cover exactly the areas where opinions start to diverge in a standards committee, and everybody has a unique solution.

Over the last years I have been asked many times how a repository of MetaLex XML files manifesting sources of law and files manifesting knowledge of those sources of law – for instance knowledge represented in LKIF – can be integrated on the technical level, and how version management of these fundamentally different kinds of things is to be aligned. In this book this integration is addressed on a conceptual level, revolving around the MetaLex notion of *representation* and the (for knowledge engineers also familiar) notion of *applicability*.

This book expands the ideas of [266, 67] in the direction of law as an institution that changes itself by legislating, and to which we often must ascribe the intention to create a certain normative order in order to benefit from our knowledge of the law. I move in these directions to explain the context in which sources of law should be

¹European Committee for Standardization

interpreted, because it is in my view necessary to do so: it is too simple to say that a proposition in a knowledge representation language represents a text fragment in a written source of law. Even when only contemplating on the domain of legislation and legislating itself, as I did for MetaLex, one cannot completely escape the issues of constitutiveness and of intentionality, because these are as relevant to legislative acts as they are to any other legal act.

Appendix B

Rechtstheorie, rechtsbronnen en het *Semantic Web*

Nederlandse samenvatting van het academisch proefschrift *Legal Theory, Sources of Law, and the Semantic Web*, te verdedigen in de Agnietenkapel van de Universiteit van Amsterdam op donderdag 25 juni 2009 door Alexander Willem Frank Boer.

B.1. Inleiding

Waar in het verleden in de kennistechnologie vooral het nut van het te bouwen kennisintensieve systeem centraal stond, verlegt de aandacht zich in toenemende mate naar de documentatie van de kennisfunctie van systemen ten behoeve van de beheersing van organisatieverandering. In de overheidsautomatisering leidt dit met name tot aandacht voor de impact van veranderend recht op de organisatie en haar kennis-systemen. Kennisrepresentatie van rechtsbronnen dient dan vooraf om de impact van veranderingen in het recht op de organisatie te analyseren, en achteraf om de werkwijze van de organisatie en haar systemen te rechtvaardigen.

Een andere recente trend is de migratie van overheidsdiensten en overheidskennis naar het Internet, in de vorm van gekoppelde *web services* en kenniscomponenten. Het *service* paradigma moedigt daarbij aan tot een toenemende formalisatie en automatisering van transacties tussen (wet-)uitvoerende organisatie en klant, en tussen uitvoerende organisaties onderling.

Deze trends hebben belangrijke consequenties voor de kennistechnologie, hier de juridische kennistechnologie in het bijzonder. In dit boek staat één van die consequenties centraal: het toenemende belang van nuttige, transparante, herbruikbare kenniscomponenten die kennis van individuele (versies van) rechtsbronnen representeren.

Dát dergelijke kenniscomponenten mogelijk zijn is een wijdverbreide en evident nuttige ontwerpaanname. De rechtsbronnen zijn immers bij uitstek de kennisbron van het recht, en de rechtsbronnen zijn het evidente middel om verandering in het recht te volgen.

De technische randvoorwaarden voor efficiënte elektronische dienstverlening over het Internet worden geleverd door een geheel van standaarden – vaak een extensie van de standaarden van het *World Wide Web* – aangeduidt als het *Semantic Web*. Andere belangrijke thema's in de zoektocht naar nuttige, transparante, herbruikbare

kenniscomponenten worden gevonden in de algemene kennistechnologie-literatuur. De drie centrale thema's in dit boek zijn:

1. het onderscheid tussen kennis van het recht en kennis van andere zaken;
2. het onderscheid tussen essentiële en contingente kennis van het recht; en
3. de toepassing van *Semantic Web* ontwerpprincipes op de representatie van rechtsbronnen.

Juridische kennistechnologie wordt gekenmerkt door een sterke nadruk op falsificeerbaarheid: ieder argument is weerlegbaar door tegen-argumenten. Dit rijmt zich slecht met zowel het gedecentraliseerde karakter van het *Semantic Web* als met de praktijk van de kennistechnologie als ontwerpdiscipline in het gemeen. Centraal staat daarom de isolatie van falsificeerbaarheid van proposities over bestaan, geldigheid, en toepasbaarheid van rechtsregels, en de vraag wat voor niet-falsificeerbare dingen – althans in die zin dat falsificatie een urgent uitvoeringsprobleem zou opleveren – we over de rechtsbronnen kunnen zeggen.

B.2. Kennisrepresentatie

Hoofdstuk 2 behandelt, na een algemene inleiding in de kennisrepresentatie en het automatisch probleem-oplossen,

1. het onderscheid tussen ontologie en epistemologie zoals gehanteerd in de relevante literatuur,
2. abstractie en ontologische stratificatie,
3. de rol van ampliatiieve (ook wel abductieve) inferenties in probleem-oplossen, en
4. het onderscheid tussen kennis en probleem-oplossen enerzijds en informatie en geformaliseerde argumentatie anderzijds.

Centraal staat een veronderstelde relatie tussen ontologische stratificatie, abstractie en ampliatiieve inferentie. Deze laatste vorm van inferentie is falsificeerbaar, in tegenstelling tot deductieve inferentie.

Een klassiek voorbeeld van abstractie in de traditie van Hobbs (cf. [150]) is de representatie van wegen als lijnen tussen punten, als vlakken, en – als men het perspectief van bijvoorbeeld Rijkswaterstaat inneemt – als een object met volume en interne structuur. Er is evident een verschil in de mate van detail van de verschillende modellen, maar een vlak wordt toch nooit een lijn door het wegnemen van details. Abstractie is dus iets heel anders dan (niet-falsificeerbare want ontologische) subsumptie.

Men past abstractie toe op een fenomeen om er grip op te krijgen. Het is dan ook de toegepaste abstractie die falsificeerbaar is, en niet als zodanig de conclusies die uit het abstracte model getrokken worden: die zijn geldig, voor het geconstrueerde model.

Ontologische stratificatie betreft het onderscheiden van verschillende niveaus van “bestaan” (cf. [254, 55]). We onderscheiden bijvoorbeeld fysieke, mentale, abstracte en sociale strata. Ook hier wordt vaak – in de superpositie van het ene object op het andere – een abstractiehierarchie aangebracht. Zo kan men bijvoorbeeld spreken van biologische, scheikundige en natuurkundige objecten, zonder te veronderstellen dat abstracte beschrijvingen van hetzelfde fenomeen in enige richting tot elkaar reduceerbaar

zijn en zelfs zonder dat identiteitscriteria op deze niveau's per se verenigbaar zijn voor dezelfde entiteit.

Het onderscheiden van ontologische strata helpt om verwarring tussen abstractie en subsumptie te vermijden, en kan ook in de representatie van het recht een nuttige rol spelen. Falsificeerbare (rechts)regels dienen dan per definitie om een abstractie naar een bepaald abstract *prototype* of exemplaar tot stand te brengen. Daarbij vindt ook een identiteitsverandering plaats.

Het volgende hoofdstuk maakt deze ideeën concreter. Hoofdstuk 3 introduceert het Semantic Web, de taal OWL DL, en de kenniscomponent. OWL DL is met uitstekend geschikt om probleemloze compositie van kenniscomponenten mogelijk te maken. De taal OWL DL dient om niet-falsificeerbare terminologische inferenties te formaliseren, en voldoet niet om falsificeerbare regels weer te geven. Ook is het in deze taal niet mogelijk om proposities over proposities van deze taal weer te geven. Omdat we wel proposities over rechtsregels moeten kunnen weergeven, is het noodzakelijk rechtsregels niet te identificeren met logische proposities van de taal.

Falsificeerbare inferentie wordt conceptueel gereduceerd tot falsificeerbare interactie met OWL DL kenniscomponenten (gebaseerd op *belief base revision*). De functie van falsificeerbare regels is dan als rechtvaardiging danwel constraint voor bepaalde belief base revision strategieën. Het gaat hier om het aanvullen van ontbrekende proposities, en dus om ampliatiieve inferentie. Omdat deze strategieën buiten het bestek van Semantic Web standaarden vallen, en bovendien betrekking hebben op algemene epistemologische kennis die niet specifiek is voor het recht, wordt geen specifieke semantiek voor deze regels uitgewerkt. Voor falsificeerbare inferentie worden twee vormen van regels beschreven als interacties met een OWL DL kenniscomponent: de *default* en de *constraint*.

In hoofdstuk 3 wordt ook een eerste voorzet gegeven voor de betekenis van het begrip *rechtsbron*. Deze zal in volgende hoofdstukken verder worden uitgewerkt; In hoofdstuk 4 wordt de rechtsinstitutie toegevoegd. De formele rechtsbron is, in de breedste zin van het woord en zoals gebruikt in MetaLex, een documentexpressie die wordt, werd, of – naar verwachting – zal worden gebruikt, om het bestaan van een bepaalde institutionele entiteit in een rechtsinstitutie te bewijzen.

B.3. Het recht als onderwerp

In hoofdstuk 4 komt het recht als domain van kennisrepresentatie aan de orde. De rechtsbron speelt in dit hoofdstuk nog nauwelijks een rol. Dit hoofdstuk behandelt de volgende onderwerpen:

1. instituties en hun constitutieve en institutionele regels, en het recht als stelsel van instituties die normen formaliseert;
2. het belang van de intentie – met name de intentie om rechtsfeiten tot stand te brengen – als voorwaarde voor de totstandkoming van een rechtsfeit;
3. normatieve regels als constitutieve regels die een negatief gewaardeerd rechtsfeit tot stand brengen;
4. De beschrijving van rechtspositie en de gerelateerde concepten (recht, plicht, bevoegdheid, etc.) in termen van een algemene epistemologische theorie voor

planning: voorkeuren, handelingsmogelijkheden, verwachtingen omtrent het gedrag van anderen, en het toeschrijven van intenties aan anderen op basis van hun gedrag.

Het recht wordt in de sociologische literatuur wel opgevat als een institutie, of stelsel van instituties (cf. [192, 248]). De institutie is een sociaal systeem met een structuur bestaande uit institutionele feiten, dat voorspelbaar interacteert met “naakte” feiten uit de omgeving van het systeem. De structuur van het systeem, en zijn interactie met de omgeving kan worden beschreven met regels: institutionele en constitutieve regels respectievelijk.

Karakteriserend voor de benadering gekozen in dit boek is dat álle (niet alleen de niet-normatieve regels of de ficties) rechtsregels uitgelegd worden in termen van de institutie, en dat bovendien onderscheid wordt gemaakt tussen “indicatieve” en noodzakelijke constitutieve voorwaarden (*requirements*) voor het totstandkomen van een rechtsfeit. Er zijn dus twee varianten constitutieve regels, en de tweede is niet-falsificeerbaar en kán dat zijn omdat de eerste dat wel is. De institutionele regel is ook niet-falsificeerbaar. Verder is kenmerkend dat voor zelfstandige institutionele entiteiten gekozen wordt, in plaats van voor gereïficeerde institutionele feiten. Dit is een innoverende toepassing van ontologische stratificatie op het recht; de indicatieve regel stelt een abstractie van een naakte handeling naar een rechtshandeling voor, waarbij de identiteit van de handeling verandert.

Terminologische inferentie, dus ook de toepassing van de institutionele regel, creeert géén nieuw rechtsfeit maar alleen een terminologische herformulering van de vastgestelde rechtsfeiten.

De rechtsinstitutie wordt gekenmerkt door doel en methode: zij probeert normatieve ordening te creëren of in stand te houden door middel van formalisatie (het opschrijven van rechtsfeiten en rechtsregels). In normale omstandigheden geldt dat 1) een rechtsfeit geconstitueerd moet worden door een onafhankelijk bestaand naakt feit, 2) het rechtsfeit uit het naakte feit afgeleid moet worden door toepassing van een rechtsregel en 3) deze rechtsregel in een rechtsbron gevonden moet kunnen worden. Als de rechtsbronnen kenbaar zijn, een beperkt corpus vormen, dan kan een rechtsfeit alleen tot stand komen door toepassing van een rechtsregel uit een specifieke set rechtsregels.

Rechtsinstituties zijn van belang omdat zij mensen redenen geven om bepaalde dingen te doen of na te laten. Zij maken gedrag voorspelbaar. Om deze functies te begrijpen moeten we een uitstap naar planning en planherkenning maken. Dit zijn algemene problemen in de kennistechnologie: er is weinig redenen hier een alternatief voor het recht voor te ontwikkelen. Bepaalde discussies in de kennisrepresentatie van het recht, zoals het *ought-to-do* versus *ought-to-be* debat, zijn dan ook alleen zinnig te bespreken in de context van een specifieke operationele theorie van planning, planherkenning, en causaliteit. Ook bepaalde concepties in de rechtstheorie, zoals Hohfeld’s rechtsrelaties, zijn alleen bruikbaar te operationaliseren binnen een dergelijke theorie.

Dit boek gaat in op handelen als het uitvoeren van een plan, en sociale rollen die men aanneemt of aan een ander toeschrijft om een geproceduraliseerde transactie met een ander uit te voeren. Omdat het *Semantic Web* opzettelijk niet in een dergelijke

universele theorie van intelligent gedrag voorziet, houdt hier ook de mogelijkheid om toepassingsneutraal te representeren al snel op.

B.4. Rechtsbronnen en hun inhoud

In hoofdstuk 5 staat de rechtsbron en de toepassingsneutrale representatie van de inhoud van de rechtsbron centraal.

De juridische kennisbron *par excellence* is de geschreven, ofwel *formele* rechtsbron. Deze rechtsbron is een bibliografische entiteit, dat wil zeggen een document of een deel daarvan. In navolging van de bibliografische FRBR standaard¹ stratificeer ik de rechtsbron *qua* item, manifestatie, expressie en werk. Het document is in eerste instantie een origineel fysiek item, dat vermenigvuldigt kan worden, resulterende in nieuwe items, die evenwel alle exemplaren zijn van dezelfde documentmanifestatie. Ook de documentmanifestatie kan bewerkt worden zonder dat de vorm van het document op informatieniveau veranderd, resulterende in nieuwe manifestaties van dezelfde documentexpressie. Toevoeging van metadata, opmaak, of omzetting van bestandsformaat zijn bijvoorbeeld dergelijke bewerkingen. Op een abstracter niveau kan het document zelfs herschreven worden, resulterend in een nieuwe expressie of editie, maar van eenzelfde werk met dezelfde bibliografische kerngegevens.

Hoofdstuk 5 geeft een compacte beschrijving van mijn werk aan de methode die gebruikt wordt in de *MetaLex* standaard (cf. [47]) om documenten en hun delen te identificeren op elk van deze niveaus.

De technische standaarden waarop het *Semantic Web* is gebaseerd zijn vooral nuttig voor identificatie en locatie van items en manifestaties. In het recht wordt vooral verwezen naar werken, en soms expressies (geconsolideerde versies). Voor kennisrepresentatie is met name het expressieniveau relevant, omdat dat het niveau waar van een kennisbron gesproken kan worden. Op dit terrein ligt nog interessant werk voor de toekomst.

Hoofdstuk 5 bespreekt onder andere verandering van rechtsbronnen op expressieniveau, en de relatie tussen het moment waarop de veranderingsgebeurtenis plaatsvindt en het toepassingmoment van een *rechtsregel* gevonden in een rechtsbron. Naast bekende fenomenen als terugwerkende kracht en uitgestelde werking wordt bijvoorbeeld ook het verschijnsel van de *ex tunc* consolidatie geduid².

De formele rechtsbron is het resultaat van een handeling die is verricht met de opzet een entiteit in een rechtsinstitutie te creëren, en dient voor derden als bewijs van het plaatsvinden van die handeling. Typerend voor de wetgevende handeling is dat tenminste een deel van die entiteiten rechtsregels zijn. Rechtsregels zijn, voorzover ze het resultaat zijn van een rechtshandeling, het resultaat van een wetgevende handeling. Door een eerder gekozen stratificatie zijn nu ook de rechtsregel en haar representatie in de bron vanzelfsprekend verschillende entiteiten.

De voornaamste functie van de rechtsbron is dus als bewijs dat bepaalde gebeurtenissen zijn voorgevallen waardoor de toestand van de rechtsinstitutie veranderd is.

¹De *Functional Requirements for Bibliographic Records* van de IFLA.

²Dit is voor Nederland nauwelijks relevant; Alleen *errata corrige* vallen in deze categorie.

De rechtsbronnen vormen dus in principe eerder een historie dan een beschrijving van het geldend recht, hoewel een goed onderhouden corpus ook die laatste functie kan vervullen. De vorm beschreven in het corpus en de vorm van de institutie kunnen ook in het meest ideale geval van elkaar afwijken, en de één kan dus ook niet zonder meer als simulatie voor de ander dienen.

In hoofdstuk 5 wordt een aantal logische proposities *over* rechtsregels onderscheiden. De werking van de *institutionele regel* en de (constitutieve) *requirement* is niet-falsificeerbaar en kan door middel van OWL axioma's weergegeven worden. De werking van indicatieve (constitutieve) regels worden weergegeven als een *default*, en die van bewijslastregels als een *constraint*.

Omdat de logische propositie geen *representatie* van de rechtsregel is, moet expliciet worden aangegeven in de regel dat de rechtsregel is toegepast om een rechtsentiteit tot stand te brengen: ook *toegepast* is een essentieel attribuut van de rechtsentiteit. Rechtsregels die de toepasbaarheid beperken (*applicability rule*; bijv. “regel *r* is alleen van toepassing op ambtenaren”) kunnen nu als *requirement* op het toegepast attribuut vormgegeven worden.

Deze benadering leidt tot een interessante observatie over zowel ficties als mede-toepasbaarheid (bijv. “regel *r* is mede toepasbaar op niet-ambtenaren”). Dit zijn nu juist de enige regels die niet goed weergegeven kunnen worden; Zij creëren namelijk in de regel een rechtsentiteit in strijd met een noodzakelijke voorwaarde (*requirement*). Het bestaan van de notie van “fictie” in het recht (terwijl je met eliminatief materialistisch gezichtspunt het hele recht als fictief zou kunnen afdoen) onderstreept dat de *requirement* wel degelijk als harde, niet-falsificeerbare regel gevoeld wordt.

Hoewel de logische propositie geen *representatie* van de rechtsregel is, is het door de gekozen logische vorm wél mogelijk geworden om de inhoud van een rechtsbron in een op zichzelf staande kenniscomponent bestaande uit logische proposities in OWL DL te representeren, zonder daarbij elementen uit andere rechtsbronnen op te hoeven nemen.

Aan de andere kant is de complicatie die ongecontroleerde falsificeerbaarheid oplevert voor het onderhoud van kennissystemen slechts vervangen door een andere: het onderscheiden van rechtsinstituties op basis van logische in plaats van inhoudelijke criteria.

B.5. Redeneren over de bedoelde normatieve orde

Hoofdstuk 6 keert terug naar de – in de rechtstheorie belangrijke – elementen van de kennisrepresentatie die minder “toepassingsneutraal” zijn. Centraal staat in het begin van het hoofdstuk de interpretatie van de normatieve regel, waarvan in hoofdstuk 4 nog werd geconstateerd dat het qua institutionele werking een “gewone” rechtsregel is, als een verplichting, verbod of permissie. De normatieve regel wordt geacht keuzes tussen alternatieve handelingen te beïnvloeden. Voor deze interpretatie wordt een weergave in OWL voorgesteld, en de adequaatheid van deze representatie wordt besproken aan de hand van verschijnselen als normatief conflict (*conflict of disaffirmation* en *compliance conflict*) en *contrary-to-duty obligations*.

Het normatief conflict geeft aanleiding to een nieuw type *applicability rule*: de *choice rule* die een keuze tussen twee anderszins gelijktijdig toepasbare regels afdwingt. Hier wordt om principiële redenen gekozen voor regels gebaseerd op beleidsregels of uitspraken van de rechter – in plaats van voor een epistemologische conflict-oplostheorie.

Het normatief conflict brengt ons ook weer terug op een aantal onderwerpen aangesneden in hoofdstuk 4. Het normatief conflict kan alleen opgevat worden als logische onverenigbaarheid van twee normatieve regels als we aan de wetgever een bedoelde normatieve orde toeschrijven. Ook de rechter is hiertoe dus genooddaakt.

Daarnaast observeert het boek dat de rechter vaak een zogenaamde *closed world assumption* moet maken over de voorhanden alternatieven in een situatie voor de handelende persoon. Omdat de rechter zich dan in feite in de kennispositie van de handelende persoon verplaatst, bestaat er feitelijk een behoorlijk grote ruimte om regels niet toepasbaar te verklaren in een bepaalde situatie.

De toegeschreven bedoelde normatieve orde is ook vaak het object van vergelijking als we aan rechtsvergelijking doen. Maar bij rechtsvergelijking kan ook het verschil tussen de bedoelde en de feitelijke normatieve ordening in een maatschappij aan de orde komen.

Uiteindelijk is in alle niet-triviale gebruikscontexten voor rechtsregels – wetgeven en -vergelijken, rechtspreken, en het alledaags handelen in interactie met andere recht-subjecten – een theorie voor praktische planning en planherkenning nodig.

Een aantal juridisch relevante onderzoeksonderwerpen zijn in het boek aangereikt, bijv. dat in die gebruikcontexten de bedoelde normatieve ordening en de kennis die aan de ander toegeschreven wordt gecontextualiseerd wordt tegen sociale rollen. Het gaat hier om een vorm van planning die noch noodzakelijke coöperatie van de omgeving (zoals in bedrijfsmatige planning) noch vanzelfsprekende tegenwerking (zoals in speltheoretische planning) veronderstelt. Dit aspect van planning in een omgeving met andere personen verdient meer aandacht in de algemene kennistechnologieliteratuur.

Bibliography

- [1] C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50:510–530, 1985.
- [2] C. E. Alchourrón and E. Bulygin. The expressive conception of norms. In R. Hilpinen, editor, *New Studies in Deontic Logic*, pages 95–125. D. Reidel Publishing Company, Dordrecht, 1981.
- [3] J. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23:123–154, 1984.
- [4] J. F. Allen and G. Ferguson. Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4(5):531–579, 1994.
- [5] L. Allen. The language of legal relations, RLL: useful in a legal ontology toolkit? In P. Visser and R. Winkels, editors, *Legal Ontologies: proceedings of the ICAIL-97 workshop*, pages 47 – 60. ACM, New York, 1997.
- [6] A. G. Amsterdam and J. S. Bruner. *Minding the Law*. Harvard University Press, Cambridge, Mass., 2000.
- [7] R. Arendsen. *Geen bericht, goed bericht: Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven*. PhD thesis, University of Amsterdam, 2008.
- [8] R. Arendsen, T. van Engers, and R. te Velde. An empirical study on business-to-government data exchange strategies to reduce the administrative costs for businesses. In R. Suomi, R. Cabral, J. Hampe, A. Heikkila, J. Jarvelainen, and E. Koskivaara, editors, *Project E-Society: Building Bricks, 6th IFIP Conference on e-Commerce, e-Business and e-Government (I3E 2006)*, volume 226 of *IFIP International Federation for Information Processing*, pages 311–323, Boston, 2006. Springer.
- [9] K. Atkinson and T. Bench-Capon. Theory and practice in ai and law: A response to branting. In M.-F. Moens and P. Spyns, editors, *Jurix 2005: The Eighteenth Annual Conference.*, Legal Knowledge and Information Systems, pages 1–13, Amsterdam, 2005. IOS Press.
- [10] J. Austin. *The Province of Jurisprudence Determined*. Weidenfeld and Nicolson, London, 1954. edited by H.L.A. Hart.
- [11] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook - Theory, Implementation and Applications*. Cambridge University Press, 2003.
- [12] F. Baader and W. Nutt. Basic description logics. In F. Baader, D. Calvanese,

- D. McGuinness, D. Nardi, and P. Patel-Schneider, editors, *Description Logic Handbook*, pages 47–100. Cambridge University Press, 2002.
- [13] F. Baader and U. Sattler. An Overview of Tableau Algorithms for Description Logics. *Studia Logica*, 69:5–40, 2001.
 - [14] J. A. Bargh. Goal \neq intent: Goal-directed thought and behavior are often unintentional. *Psychological Inquiry*, 1(3):248–251, 1990.
 - [15] J. Barwise and J. Perry. *Situations and Attitudes*. Bradford Books, MIT Press, 1983.
 - [16] T. Bench-Capon. Neural networks and open texture. In *ICAIL '93: Proceedings of the 4th international conference on Artificial intelligence and law*, pages 292–297, New York, NY, USA, 1993. ACM.
 - [17] T. Bench-Capon and F. Coenen. Exploiting isomorphism: development of a KBS to support British coal insurance claims. In M. Sergot, editor, *Proceedings of the third International Conference on AI and Law*, pages 62–69, New York, 1991. Acm.
 - [18] T. Bench-Capon and G. Sartor. Using values and theories to resolve disagreement in law. In J. Breuker, R. Leenes, and R. Winkels, editors, *Jurix 2000: The Thirteenth Annual Conference.*, Legal Knowledge and Information Systems, pages 73–84, Amsterdam, 2000. IOS Press.
 - [19] T. Bench-Capon and G. Sartor. A model of legal reasoning with cases incorporating theories and values. *Artificial Intelligence*, 150:97–143, 2003.
 - [20] T. J. M. Bench-Capon. Deep models, normative reasoning and legal expert systems. In *ICAIL '89: Proceedings of the 2nd international conference on Artificial intelligence and law*, pages 37–45, New York, NY, USA, 1989. ACM.
 - [21] T. J. M. Bench-Capon, G. O. Robinson, T. W. Routen, and M. J. Sergot. Logic programming for large scale applications in law: A formalisation of supplementary benefit legislation. In *ICAIL '87: Proceedings of the 1st international conference on Artificial intelligence and law*, pages 190–198, New York, NY, USA, 1987. ACM.
 - [22] V. R. Benjamins, P. Casanovas, J. Breuker, and A. Gangemi, editors. *Law and the Semantic Web*. Springer, LNCS 3369, 2005.
 - [23] F. Bex, H. Prakken, and B. Verheij. Formalising argumentative story-based analysis of evidence. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Law*. New York, ACM Press, 2007.
 - [24] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
 - [25] J. Blythe. Decision-theoretic planning. *AI Magazine*, 20(2), 1999.
 - [26] O. Bodenreider, B. Smith, and A. Burgun. The ontology-epistemology divide: a case study in medical terminology. In A. Varzi and L. Vieu, editors, *Formal ontology in Information Systems*, pages 185–198. IOS-Press, Amsterdam, 2004.
 - [27] G. Boella, L. Lesmo, and R. Damiano. On the ontological status of norms. In *Law and the Semantic Web*, volume 3369 of *Lecture Notes in Computer Science*. Springer Verlag, 2005.
 - [28] G. Boella and L. van der Torre. Rational norm creation: attributing mental attitudes to normative systems, part 2. In *ICAIL '03: Proceedings of the 9th*

- international conference on Artificial intelligence and law*, pages 81–82, New York, NY, USA, 2003. ACM.
- [29] G. Boella and L. W. N. van der Torre. Regulative and constitutive norms in normative multiagent systems. In *Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning*, Whistler (CA), 2004.
 - [30] A. Boer. The Consultancy Game. In J. Breuker, R. Leenes, and R. Winkels, editors, *Legal Knowledge and Information Systems. Jurix 2000: The Thirteenth Annual Conference*, Frontiers in Artificial Intelligence and Applications, pages 99–112, Amsterdam, 2000. IOS Press.
 - [31] A. Boer. Notes on the legal knowledge interchange format. Estrella technical report, Leibniz Center for Law, University of Amsterdam, 2006.
 - [32] A. Boer. Using event descriptions for metadata about legal documents. In R. Winkels and E. Francesconi, editors, *Electronic Proceedings of the Workshop on Standards for Legislative XML, in conjunction with Jurix 2007*, 2007.
 - [33] A. Boer and J. Breuker. So you want to Validate your Problem-Solving Methods? In *Proceedings of the Knowledge Acquisition Workshop (KAW)*, Banff, Canada, 1998.
 - [34] A. Boer and J. Breuker. Verifying and Validating the CommonKADS Library of Problem-Solving Methods. In *Proceedings of the Second French-American Conference on Validation and Verification*, Nîmes, France, 1998.
 - [35] A. Boer, T. F. Gordon, K. van den Berg, M. Di Bello, A. Föhrécz, and R. Vas. Specification of the legal knowledge interchange format. Deliverable 1.1, Estrella, 2007.
 - [36] A. Boer, R. Hoekstra, R. Winkels, and T. van Engers. *METAlex*: Jurisdiction and Language. In M. Palmirani, T. van Engers, and M. A. Wimmer, editors, *Proceedings of the E-Government Workshop in conjunction with JURIX 2003*, pages 54–66. Universitätsverlag Rudolf Trauner, December 2003.
 - [37] A. Boer, R. Hoekstra, R. Winkels, T. van Engers, and F. Willaert. *METAlex*: Legislation in XML. In T. Bench-Capon, A. Daskalopulu, and R. Winkels, editors, *Legal Knowledge and Information Systems. Jurix 2002: The Fifteenth Annual Conference*, Frontiers in Artificial Intelligence and Applications, pages 1–10, Amsterdam, 2002. IOS Press.
 - [38] A. Boer, R. Hoekstra, R. Winkels, T. van Engers, and F. Willaert. Proposal for a Dutch Legal XML Standard. In R. Traunmüller and K. Lenk, editors, *Electronic Government (EGOV 2002)*, pages 142–149, Berlin; Heidelberg; New York, 2002. Springer Lecture Notes in Computer Science.
 - [39] A. Boer, E. Hupkes, F. Vitali, M. Palmirani, and B. Rátai. Metalex CEN workshop proposal. Technical report, CEN Workshop on an Open XML Interchange Format for Legal and Legislative Resources (Metalex), May 2008.
 - [40] A. Boer and T. van Engers. Standards for spatial regulations. In *Proceedings of the Workshop on the Role of legal knowledge in e-Government, held in conjunction with the Tenth International Conference on ARTIFICIAL INTELLIGENCE and LAW*, Bologna, Italy, 2005.
 - [41] A. Boer, T. van Engers, R. Peters, and R. Winkels. Separating law from geograp-

- hy in gis-based egovernment services. *Artificial Intelligence & Law*, 15(1):49–76, February 2007.
- [42] A. Boer, T. van Engers, and R. Winkels. Using Ontologies for Comparing and Harmonizing Legislation. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, Edinburgh (UK), 2003. ACM Press.
 - [43] A. Boer, T. van Engers, and R. Winkels. Mixing legal and non-legal norms. In M.-F. Moens and P. Spyns, editors, *Legal Knowledge and Information Systems. Jurix 2005: The Eighteenth Annual Conference*, Frontiers in Artificial Intelligence and Applications, pages 25–36, Amsterdam, 2005. IOS Press.
 - [44] A. Boer, T. van Engers, and R. Winkels. Normative statements on the semantic web. In *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques, held in conjunction with the Tenth International Conference on ARTIFICIAL INTELLIGENCE and LAW*, Bologna, Italy, 2005.
 - [45] A. Boer, T. van Engers, and R. Winkels. Open standards for spatial regulations: an interdisciplinary approach. In *Proceedings of the Holland Open Software Conference*, Amsterdam, Netherlands, 2005.
 - [46] A. Boer, T. van Engers, and R. Winkels. Probleemverkenning implementatie afspraken en standaarden DURP. Rapport voor ministerie VROM, Leibniz Center for Law, April 2005.
 - [47] A. Boer, F. Vitali, and E. de Maat. CEN Workshop Agreement on MetaLex XML, an open XML Interchange Format for Legal and Legislative Resources (CWA 15710). Technical report, European Committee for Standardization (CEN), 2006.
 - [48] A. Boer and R. Winkels. What’s in an Interchange Standard for Legislative XML? *I Quaderni*, Novembre(18):32–41, 2005.
 - [49] A. Boer, R. Winkels, and R. Hoekstra. *METAlex*: een schema voor de structuur van wet- en regelgeving. *<!ELEMENT*, 1, 2003.
 - [50] A. Boer, R. Winkels, R. Hoekstra, and T. M. van Engers. Knowledge Management for Legislative Drafting in an International Setting. In D. Bourcier, editor, *Legal Knowledge and Information Systems. Jurix 2003: The Sixteenth Annual Conference.*, Frontiers in Artificial Intelligence and Applications, pages 91–100, Amsterdam, 2003. IOS Press.
 - [51] A. Boer, R. Winkels, T. van Engers, and E. de Maat. A content management system based on an event-based model of version management information in legislation. In T. Gordon, editor, *Legal Knowledge and Information Systems. Jurix 2004: The Seventeenth Annual Conference.*, Frontiers in Artificial Intelligence and Applications, pages 19–28, Amsterdam, 2004. IOS Press.
 - [52] A. Boer, R. Winkels, T. van Engers, and E. de Maat. Time and versions in *METAlex* XML. In *Proceeding of the Workshop on Legislative XML*, Kobaek Strand, 2004.
 - [53] A. Boer, R. Winkels, and F. Vitali. Proposed XML standards for law: Metalex and LKIF. In A. R. Lodder and L. Mommers, editors, *Legal Knowledge and Information Systems. Jurix 2007: The Twentieth Annual Conference Annual Conference*, volume 165 of *Frontiers in Artificial Intelligence and Applications*, pages 19–28. IOS Press, December 2007.

- [54] A. Boer, R. Winkels, and F. Vitali. Metalex XML and the Legal Knowledge Interchange Format. In G. Sartor, P. Casanovas, N. Casellas, and R. Rubino, editors, *Computational Models of the Law*, volume LNCS 4884 of *Lecture Notes in Artificial Intelligence*. Springer, 2008.
- [55] S. Borgo, N. Guarino, and C. Masolo. Stratified ontologies: The case of physical objects. In *Proceedings of the ECAI-96 Workshop on Ontological Engineering*, 1996.
- [56] C. Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68, 1994.
- [57] I. Bratko. *Prolog programming for artificial intelligence, 3rd ed.* Harlow: Addison-Wesley, 2001.
- [58] M. Bratman. Two faces of intention. *Philosophical Review*, 93(3):375–405, 1984.
- [59] J. Breuker. A suite of problem types. In J. Breuker and W. V. de Velde, editors, *CommonKADS Library for Expertise Modelling*, pages 57 – 88. IOS-Press/Ohmsha, Amsterdam/Tokyo, 1994.
- [60] J. Breuker. The construction and use of ontologies of criminal law in the eCourt european project. In V. Golaczynski, editor, *Proceedings of Means of electronic communication in court administration*, pages 15–40. ILV, Wroclaw, 2003.
- [61] J. Breuker, A. Boer, R. Hoekstra, and K. van den Berg. Developing content for LKIF: Ontologies and frameworks for legal reasoning. In T. M. van Engers, editor, *Legal Knowledge and Information Systems. Jurix 2006: The Nineteenth Annual Conference*, volume 152 of *Frontiers in Artificial Intelligence and Applications*, pages 169–174. IOS Press, December 2006.
- [62] J. Breuker and W. V. de Velde, editors. *CommonKADS Library for Expertise Modeling: reusable problem solving components*. IOS-Press/Ohmsha, Amsterdam/Tokyo, 1994.
- [63] J. Breuker and N. Den Haan. Separating world and regulation knowledge: where is the logic? In *Proceedings of the third international conference on AI and Law*, Oxford, 1991. Acm.
- [64] J. Breuker, A. Elhag, E. Petkov, and R. Winkels. It support for the judiciary: Use of ontologies in the eCourt project. In *Proceedings of the ICCS Conference on Conceptual Graphs*, pages 17 – 26, 2002.
- [65] J. Breuker and R. Hoekstra. Core concepts of law: taking common-sense seriously. In *Proceedings of Formal Ontologies in Information Systems FOIS-2004*, pages 210–221. IOS-Press, 2004.
- [66] J. Breuker, R. Hoekstra, A. Boer, K. van den Berg, R. Rubino, G. Sartor, M. Palmirani, A. Wyner, and T. Bench-Capon. OWL ontology of basic legal concepts (LKIF-Core). Deliverable 1.4, Estrella, 2007.
- [67] J. Breuker and R. Winkels. Taking ontological commitments seriously. *Extraction Des Connaissances Et Apprentissage*, 1(4):9–26, 2002.
- [68] R. Brohm. Bringing Polanyi on the Theatre Stage: A study on Polanyi applied to Knowledge Management. In J. F. Schreinemakers and T. V. Engers, editors, *15 Years of Knowledge Management*, volume 3 of *Advances in Knowledge Management*, chapter Part I, pages 3–24. Ergon Verlag, 2007.
- [69] E. Bulygin. On norms of competence. *Law and Philosophy*, 11(3):201–216, 1992.

- [70] M. Bussani and U. Mattei, editors. *Making European Law: Essays on the Common Core Project*. Università degli Studi di Trento, 2000.
- [71] J. Caldw. Cinderella cities. See <http://www-1.ibm.com/industries-/government/ieg/pdf/cinderella.pdf>, 2000.
- [72] D. Calvanese, G. D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Epistemic first-order queries over description logic knowledge bases. In *Description Logics*, 2006.
- [73] D. Calvanese, G. D. Giacomo, and M. Lenzerini. A Framework for Ontology Integration. In *Proceedings of the First Semantic Web Working Symposium*, pages 303–316, 2001.
- [74] H.-N. Castañeda. On the semantics of the ought-to-do. *Synthese*, 21(3-4):449–468, 1970.
- [75] B. Chandrasekaran. Design problem solving: A task analysis. *AI Magazine*, 11(4):59–71, 1990.
- [76] B. Chandrasekaran and T. R. Johnson. Generic tasks and task structures: History, critique and new directions. In J. M. David, J. P. Krivine, and R. Simmons, editors, *Second Generation Expert Systems*. Springer Verlag, 1993.
- [77] R. M. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24:33–36, 1963.
- [78] R. M. Chisholm. Legal reasoning in 3-d. In A. Press, editor, *In Proceedings of the First International Conference in Artificial Intelligence and Law*, pages 155–163, Boston, 1987.
- [79] A. Chorley and T. Bench-Capon. Reasoning with legal cases as theory construction: Some experimental results. In D. Bourcier, editor, *Jurix 2003: The Sixteenth Annual Conference.*, Legal Knowledge and Information Systems, pages 173–182, Amsterdam, 2003. IOS Press.
- [80] A. Chorley and T. Bench-Capon. Agatha: Automation of the construction of theories in case law domains. In T. Gordon, editor, *Jurix 2004: The Seventeenth Annual Conference.*, Legal Knowledge and Information Systems, pages 89–98, Amsterdam, 2004. IOS Press.
- [81] W. J. Clancey. Heuristic classification. *Artificial Intelligence*, 27:289–350, 1985.
- [82] W. J. Clancey. Model construction operators. *Artificial Intelligence*, 53(1):1–115, 1992.
- [83] S. Cnossen. Vermogensrendementsheffing: vondst of miskleun? OCFEB papers and proceedings 9903, Onderzoekscentrum Financieel Economisch Beleid (OCFEB), Erasmus Universiteit Rotterdam, 1999.
- [84] J. Cohen and I. Stewart. *The Collapse of Chaos: Discovering Simplicity in a Complex World*. Penguin Books, 1994.
- [85] P. Cohen and H. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.
- [86] T. Costello. Change, change, change: Three approaches. In *IJCAI*, pages 1426–1433, 1997.
- [87] P. d’Altan, J.-J. C. Meyer, and R. Wieringa. An integrated framework for ought-to-be and ought-to-do constraints. *Artif. Intell. Law*, 4(2):77–111, 1996.
- [88] R. Davis, H. E. Shrobe, and P. Szlovits. What is knowledge representation? *AI*

- Magazine*, 14(1):17–33, 1993.
- [89] E. Dayton. Two approaches to deontic logic. *The Journal of Value Inquiry*, 15(2):137–147, 1981.
 - [90] M. de Montaigne. Of ancient customs. In *Essays of Michel de Montaigne*. Project Gutenberg (online), 2006/1580.
 - [91] C. de Montesquieu. *Spirit of the Laws*. Cambridge University Press, 1989.
 - [92] N. den Haan. On structure preserving representations of law. In T. Gordon, editor, *Proceedings of the German Conference on Artificial Intelligence and Law*, Vienna, 1994. Invited lecture.
 - [93] N. den Haan and J. Breuker. Constructing normative rules. In R. van Kralingen, H. van den Herik, J. Prins, M. Sergot, and J. Zeleznikow, editors, *Legal Knowledge Based Systems JURIX'96: Foundations of legal knowledge systems*, pages 135–147. Tilburg University Press, 1996. ISBN 90-361-9657-4.
 - [94] N. den Haan and R. Winkels. The Deep Structure of Law. In H. Prakken, A. Muntjewerff, A. Soeteman, and R. Winkels, editors, *Legal Knowledge Based Systems: The Relation with Legal Theory*, pages 43–54. Koninklijke Vermande, Lelystad, 1994.
 - [95] D. Dennett. *The Intentional Stance*. MIT-Press, 1987.
 - [96] J. P. Dick. Representation of legal text for conceptual retrieval. In *ICAIL '91: Proceedings of the 3rd international conference on Artificial intelligence and law*, pages 244–253, New York, NY, USA, 1991. ACM Press.
 - [97] F. Dignum, J. Broersen, V. Dignum, and J.-J. Meyer. Meeting the deadline: Why, when and how. In *Formal Approaches to Agent-Based Systems*. Lecture Notes in Computer Science, Springer Verlag, 2005.
 - [98] C. Doctorow. Metacrap: Putting the torch to seven straw-men of the meta-utopia. See <http://www.well.com/doctorow/metacrap.htm>, 2001.
 - [99] M. Donnelly. Relative places. *Applied Ontology*, 1(1):55–75, 2005.
 - [100] D. Doyle. Prospects for preferences. *Computational Intelligence*, 20:111–136, 2004.
 - [101] J. Doyle and R. H. Thomason, editors. *Working papers of the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning*. American Association for Artificial Intelligence, Menlo Park, California, 1997.
 - [102] J. Doyle and M. P. Wellman. Representing preferences as ceteris paribus comparatives. In *Decision-Theoretic Planning: Papers from the 1994 Spring AAAI Symposium*, pages 69–75. AAAI Press, Menlo Park, California, 1994.
 - [103] W. Dubislav. Zur unbegründbarkeit der forderungssätze. *Theoria*, 3:330–342, 1937.
 - [104] J. Dyer. MAUT multiattribute utility theory. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 265–285. Springer Verlag, Boston, Dordrecht, London, 2005.
 - [105] A. A. Elhag, J. A. Breuker, and B. W. Brouwer. On the formal analysis of normative conflicts. In H. van den Herik et al., editor, *JURIX 1999: The Twelfth Annual Conference*, Frontiers in Artificial Intelligence and Applications, pages 35–46, Nijmegen, 1999. GNI.
 - [106] E. Engelen. Financialization, Pension Restructuring, and the Logic of Funding.

- In *Proceedings of SASE 2002*, 2002.
- [107] K. Erol, J. Hendler, and D. S. Nau. Htn planning: complexity and expressivity. In *AAAI'94: Proceedings of the twelfth national conference on Artificial intelligence (vol. 2)*, pages 1123–1128, Menlo Park, CA, USA, 1994. American Association for Artificial Intelligence.
 - [108] D. Fensel. A case study: Assumptions and limitations of a problem solving method. In *International Journal of Human Computer Studies*, 1995.
 - [109] D. Fensel and R. Straatman. Problem-solving methods: Making assumptions for efficiency reasons. In *EKAW '96: Proceedings of the 9th European Knowledge Acquisition Workshop on Advances in Knowledge Acquisition*, pages 17–32, London, UK, 1996. Springer-Verlag.
 - [110] K. M. Ford, F. E. Petry, J. R. Adams-Webber, and P. J. Chang. An approach to knowledge acquisition based on the structure of personal construct systems. *IEEE Transactions on Knowledge and Data Engineering*, 3(1):78–88, 1991.
 - [111] J. W. Forrester. Gentle murder, or the adverbial samaritan. *Journal of Philosophy*, 81:193–197, 1984.
 - [112] E. Francesconi. *Technologies for European Integration*. European Press Academic Publishing, 2007.
 - [113] B. Gaines and M. Shaw. Knowledge acquisition tools based on personal construct psychology, 1993.
 - [114] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening ontologies with DOLCE. In A. Gomez-Perez and V. Benjamins, editors, *Proceedings of the EKAW-2002*, pages 166–181. Springer, 2002.
 - [115] A. Gangemi, D. Pisanelli, and G. Steve. A formal ontology framework to represent norm dynamics. In R. Winkels and R. Hoekstra, editors, *Proceedings of the Second International Workshop on Legal Ontologies (LEGONT)*, Amsterdam, Netherlands, 2001.
 - [116] A. Gangemi, A. Prisco, M. Sagri, G. Steve, and D. Tiscornia. Some ontological tools to support legal regulatory compliance, with a case study. In *Proceedings of the Workshop on Regulatory Ontologies and the Modeling of Complaint Regulations (WORM CoRe 2003)*, Catania, Sicily, Italy, 2003. Springer Lecture Notes in Computer Science.
 - [117] P. Gärdenfors and D. Makinson. Revisions of knowledge systems using epistemic entrenchment. In M. Y. Vardi, editor, *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the Second Conference*, pages 83–96. Morgan Kaufmann, Los Altos, CA, 1988.
 - [118] S. Garrels. Imitation, mirror neurons, and mimetic desire: Convergence between the mimetic theory of René Girard and empirical research on imitation. *Contagion: Journal of Violence, Mimesis, and Culture*, 12–13:47–86, 2006.
 - [119] T. Geiger. *Theodor Geiger on Social Order and Mass Society: Selected Papers*. University of Chicago Press, 1969.
 - [120] M. R. Genesereth and N. J. Nilsson. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann, Los Altos, 1987.
 - [121] A. I. Goldman. Desire, intention, and the simulation theory. In B. F. Malle, L. J. Moses, and D. A. Baldwin, editors, *Intentions and intentionality: Foundations*

- of social cognition*, pages 207–225. MIT Press, Cambridge, MA, 2001.
- [122] T. F. Gordon, H. Prakken, and D. Walton. The Carneades model of argument and burden of proof. *Artificial Intelligence*, 171(10-11):875–896, 2007.
 - [123] G. Governatori and A. Rotolo. Modelling contract using RuleML. In T. Gordon, editor, *Legal Knowledge and Information Systems (Jurix 2004)*, pages 141–150, Amsterdam, 2004. IOS Press.
 - [124] T. Grandon Gill. Early expert systems: Where are they now? *MIS Quarterly*, 19(1):51–81, 1995.
 - [125] B. C. Grau, I. Horrocks, and B. Parsia. Online proceedings of the first OWL: Experiences and directions workshop, 2005. <http://www.mindswap.org/2005/OWLWorkshop/accepted.shtml>.
 - [126] B. C. Grau, B. Parsia, E. Sirin, and A. Kalyanpur. Modularizing OWL ontologies. In *Proc. of the 4th International Semantic Web Conference (ISWC-2005)*, 2005.
 - [127] S. Grimm and B. Motik. Closed world reasoning in the semantic web through epistemic operators. In *Online Proceedings of the First OWL: Experiences and Directions Workshop (OWLED)*, Galway, Ireland, 2005. <http://dip.semanticweb.org/publications.html>.
 - [128] C. Groendijk. Neural schemata in automated judicial problem solving. In C. Grütters, J. Breuker, H. van den Herik, A. Schmidt, and C. de Vey Mestdagh, editors, *Legal Knowledge Based Systems: Information Technology & Law*, Lelystad, 1992. Vermande.
 - [129] M. M. Groothuis and J. S. Svensson. Expert system support and juridical quality. In J. Breuker, R. Leenes, and R. Winkels, editors, *Legal Knowledge and Information Systems. Jurix 2000: The Thirteenth Annual Conference*, Frontiers in Artificial Intelligence and Applications, pages 1–10, Amsterdam, 2000. IOS Press.
 - [130] D. Grossi and F. Dignum. Abstract and concrete norms in institutions. Technical report, University of Utrecht, technical report UU-CS-2004-016, 2004.
 - [131] T. R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.
 - [132] J. Hage. De wondere wereld van het recht. Technical report, Universiteit Maastricht, 2007. inaugurele rede.
 - [133] J. C. Hage, H. B. Verheij, and A. R. Lodder. Reason based logic; a logic that deals with rules and reasons. In *In J.M. Akkermans and J.A. Breuker, editors, Working Papers NAIC '93*, pages 293–304, 1993.
 - [134] C. Halaschek-Wiener, Y. Katz, and B. Parsia. Belief base revision for expressive description logics. In *Online Proceedings of the Second OWL: Experiences and Directions Workshop (OWLED)*, Athens, Georgia, USA, 2006. <http://owl-workshop.man.ac.uk/accepted06.shtml>.
 - [135] A. Halpin. Fundamental legal conceptions reconsidered. *Canadian Journal of Law and Jurisprudence*, 16:41–54, 2003.
 - [136] B. Hansson. An analysis of some deontic logics. *Nous*, 3:373–398, 1969.
 - [137] S. O. Hansson. Decision theory: A brief introduction. Technical report, Department of Philosophy and the History of Technology, Royal Institute of Technology,

- Stockholm, 1994.
- [138] S. O. Hansson. *The Structure of Values and Norms*. Cambridge University Press, UK, 2001.
 - [139] S. O. Hansson. The dual nature of technological concepts. In *Invited contribution to the international conference Technological Knowledge: Philosophical Reflection*, 2002.
 - [140] H. Hart. *The Concept of Law*. Clarendon Press, Oxford, 1961.
 - [141] H. Hart and T. Honoré. *Causation in the Law*. Oxford University Press, New York, second edition edition, 1985.
 - [142] P. Hassett. Can expert systems improve the exercise of judicial discretion? In C. Grütters, J. Breuker, H. van den Herik, A. Schmidt, and C. de Vey Mestdag, editors, *Legal Knowledge Based Systems: Information Technology and Law*, 1992.
 - [143] P. J. Hayes. Naive physics I: Ontology for liquids. In J. R. Hobbs and R. C. Moore, editors, *Formal Theories of the Common Sense World*, pages 71–108. Ablex Publishing Corporation, Norwood, 1985.
 - [144] P. J. Hayes. The second naive physics manifesto. In J. R. Hobbs and R. C. Moore, editors, *Formal Theories of the Common Sense World*, pages 1–36. Ablex Publishing Corporation, Norwood, 1985.
 - [145] W. Hettich and S. L. Winer. Economic and political foundations of tax structure. *American Economic Review*, 78(4):701–712, 1988.
 - [146] W. Hettich and S. L. Winer. Rules, Politics, and the Normative Analysis of Taxation. Carleton Economic Papers 00-12, Carleton University, 2000.
 - [147] H. Hill. A functional taxonomy of normative conflict. *Law and Philosophy*, 6:227–247, 1987.
 - [148] F. A. Hindriks. *Rules & Institutions; essays in meaning, speech and social ontology*. PhD thesis, Erasmus University Rotterdam, 2005.
 - [149] J. Hobbs and V. Kreinovich. Optimal choice of granularity in commonsense estimation: Why half-orders of magnitude. In *Proc. IFSA/NAFIPS'2001*, pages 1343–1348, 2001.
 - [150] J. R. Hobbs. Sketch of an ontology underlying the way we talk about the world. *Int. J. Hum.-Comput. Stud.*, 43(5-6):819–830, 1995.
 - [151] J. R. Hobbs and F. Pan. Time ontology in OWL. Technical report, W3C, May 2006.
 - [152] R. Hoekstra and J. Breuker. Commonsense causal explanation in a legal domain. *Artificial Intelligence & Law*, 15(3):281–299, September 2007. Special issue on Legal Ontologies and Artificial Intelligence Techniques.
 - [153] R. Hoekstra, J. Breuker, M. D. Bello, and A. Boer. The LKIF Core ontology of basic legal concepts. In P. Casanovas, M. A. Biasiotti, E. Francesconi, and M. T. Sagri, editors, *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, June 2007.
 - [154] R. Hoekstra, J. Breuker, M. Di Bello, and A. Boer. LKIF Core: Principled ontology development for the legal domain. In J. Breuker, P. Casanovas, M. Klein, and E. Francesconi, editors, *Law, Ontologies, and the Semantic Web*. IOS Press, 2008. Submitted for review.
 - [155] R. Hoekstra, J. Breuker, M. Di Bello, and A. Boer. LKIF core: Principled onto-

- logy development for the legal domain. In J. Breuker, P. Casanovas, M. Klein, and E. Francesconi, editors, *Law, Ontologies and the Semantic Web*, Frontiers of Artificial Intelligence and Applications. IOS Press, Amsterdam, 2008. To be Published.
- [156] W. Hohfeld. *Fundamental Legal Conceptions as Applied in Legal Reasoning*. Yale University Press, 1919. Edited by W.W. Cook, fourth printing, 1966.
 - [157] I. Horrocks and P. Patel-Schneider. Reducing OWL entailment to description logic satisfiability. In *Proc. of the 2nd International Semantic Web Conference (ISWC)*, 2003.
 - [158] I. Horrocks, P. Patel-Schneider, and F. van Harmelen. From SHIQ and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26, 2003.
 - [159] I. Horrocks and S. Tessaris. A conjunctive query language for description logic aboxes. In *AAAI/IAAI*, pages 399–404, 2000.
 - [160] D. Hume. *A Treatise on Human Nature*. Boston: Little, Brown and Company, 1854 (1739).
 - [161] D. Hunter. Out of their minds: Legal theory in neural networks. *Artificial Intelligence and Law*, 7:129–151, 1999.
 - [162] A. Jones and M. Sergot. *On the Characterization of Law and Computer Systems: The Normative Systems Perspective*, chapter 12, pages 275–307. John Wiley & Sons, 1993.
 - [163] J. Jørgensen. Imperatives and logic. *Erkenntnis*, 7:288–296, 1937.
 - [164] D. Kahneman, J. Knetsch, and R. Thaler. Fairness as a constraint on profit-seeking. *American Economic Review*, 76(4):728–741, 1986.
 - [165] D. Kahneman, J. Knetsch, and R. Thaler. The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1):193–206, 1991.
 - [166] D. Kahneman and A. Tversky. Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.
 - [167] A. Katsenelinboigen. *The Concept of Indeterminism and Its Applications: Economics, Social Systems, Ethics, Artificial Intelligence, and Aesthetics*. Greenwood Publishing, Westport, 1997.
 - [168] Y. Katz and J. Golbeck. Social network-based trust in prioritized default logic. In *AAAI*. AAAI Press, 2006.
 - [169] R. Keeney and H. Raiffa. *Decisions with multiple objectives*. Wiley Press, 1976.
 - [170] M. Keller, M. Gummeran, X. Wang, and S. Lindsey. Understanding perspectives and emotions in contract violation: Development of deontic and moral reasoning. *Child Development*, 75:614–635, 2004.
 - [171] G. A. Kelly. *The Psychology of Personal Constructs*. Norton, New York, 1955.
 - [172] H. Kelsen. *General Theory of Norms*. Clarendon Press, Oxford, 1991.
 - [173] M. Klein and N. F. Noy. A component-based framework for ontology evolution. Technical Report IR-504, Department of Computer Science, Vrije Universiteit Amsterdam, March 2003.
 - [174] M. Kohen. Europe and the Standardization of the Law: Past and Present. In *Proceedings of the Conference on Globalisation and International Relations in the 21st Century*, 2002.

- [175] V. Kolovski, B. Parsia, and Y. Katz. Implementing OWL defaults. Technical report, Defense Technical Information Center OAI-PMH Repository [<http://stinet.dtic.mil/oai/oai>] (United States), 2006.
- [176] T. Konrad. On law, religion, and custom. In *To Honor René Girard: Presented on the Occasion of his Sixtieth Birthday by Colleagues, Students, Friends*, pages 177–190, Saratoga, 1986.
- [177] W. Kowalczyk. Neural networks in knowledge-based systems. In *Proceedings of the second symposium on neural networks*, Nijmegen, 1992.
- [178] R. A. Kowalski and M. J. Sergot. A logic-based calculus of events. *New Generation Comput.*, 4(1):67–95, 1986.
- [179] R. A. Kowalski and F. Toni. Abstract argumentation. *Artificial Intelligence and Law*, 4:275–296, 1996.
- [180] M. Krötzsch, S. Rudolph, and P. Hitzler. Expressive tractable description logics based on SROIQ rules. Technical report, Universität Karlsruhe (TH), February 2008.
- [181] J. Kuhl. Who controls whom when i control myself? *Psychological Inquiry*, 7(1):61–68, 1996.
- [182] C. Lagoze, J. Hunter, and D. Brickley. An event-aware model for metadata interoperability. In *ECDL 2000 - Lisbon, September 2000*, 2000.
- [183] G. Lakoff and M. Johnson. *Metaphors We Live By*. University of Chicago Press, Chicago, 1980.
- [184] J. Lehmann. *Causation in Artificial Intelligence and Law - A modelling approach*. PhD thesis, University of Amsterdam - Faculty of Law - Department of Computer Science and Law, 2003.
- [185] D. Lenat and R. Guha. *Building Large Knowledge Based Systems*. Addison Wesley, Reading, Massachusetts, 1990.
- [186] D. B. Lenat. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.
- [187] H. J. Levesque. Foundations of a functional approach to knowledge representation. *Artif. Intell.*, 23(2):155–212, 1984.
- [188] S. Levinson. *Pragmatics*. Cambridge University Press, Cambridge, 1983.
- [189] L. Lindahl. Conflicts in systems of legal norms: A logical point of view. In P. Brouwer, B. Hol, A. Soeteman, W. van der Velde, and A. de Wild, editors, *Coherence and Conflict in Law*, pages 39–64, Deventer, 1992. Kluwer.
- [190] C. Lupo, F. Vitali, E. Francesconi, M. Palmirani, R. Winkels, E. de Maat, A. Boer, and P. Mascellani. General XML format(s) for legal sources. Deliverable 3.1, Estrella, 2007.
- [191] C. Maasdorp. Concept and Context: Tacit Knowledge in Knowledge Management Theory. In J. F. Schreinemakers and T. V. Engers, editors, *15 Years of Knowledge Management*, volume 3 of *Advances in Knowledge Management*, chapter Part I, pages 59–68. Ergon Verlag, 2007.
- [192] N. MacCormick. Norms, institutions, and institutional facts. *Law and Philosophy*, 17(3):301–345, 1998.
- [193] R. M. MacGregor and D. L. McGuinness. DARPA’s high performance knowledge base (HPKB) program. In *Description Logics*, 1999.

- [194] D. Makinson. Five faces of minimality. *Studia Logica*, 52(3):339–380, 1993.
- [195] D. Makinson. On a fundamental problem of deontic logic. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems. New Studies in Deontic Logic and Computer Science*, volume 49 of *Frontiers in Artificial Intelligence and Applications*, pages 29–53. IOS Press, Amsterdam, 1999.
- [196] R. H. Marin, J. L. Hernandez, and J. J. I. Delgado. Cronolex: a computing representation of law dynamics. In *XXII World congress of Philosophy of Law and Socialphilosophy. Granada, Spain*, May 2005.
- [197] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. Wonder-Web ontology library. Technical Report Deliverable D18, version 1, ISTC-CNR (Italy), 2002.
- [198] C. Masolo, L. Vieu, E. Bottazzi, C. Catenacci, R. Ferrario, A. Gangemi, and N. Guarino. Social roles and their descriptions. In *Proceedings of Knowledge Representation Workshop*, 2004.
- [199] T. Mazzarese. Towards the semantics of “constitutive” in judicial reasoning. *Ratio Iuris*, 12:252–262, 1999.
- [200] J. McCarthy. Programs with common sense. In *Proceedings of the Teddington Conference on the Mechanization of Thought Processes*, pages 75–91, London, 1959. Her Majesty’s Stationary Office.
- [201] R. N. Moles. Expert systems - the need for theory. In C. Grütters, J. Breuker, H. van den Herik, A. Schmidt, and C. de Vey Mestdagh, editors, *Legal Knowledge Based Systems: Information Technology and Law*, 1992.
- [202] L. Mommers. *Applied Legal Epistemology*. PhD thesis, Universiteit Leiden, 2002.
- [203] J. Montrose. Is negligence an ethical or a sociological concept? *Modern Law Review*, 21(3):259–264, 1958.
- [204] M. S. Moore. Legal reality: a naturalist approach to legal ontology. *Law and Philosophy*, 21:619–705, 2002.
- [205] B. Motik, I. Horrocks, R. Rosati, and U. Sattler. Can OWL and logic programming live together happily ever after? In I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo, editors, *The Semantic Web - ISWC 2006*, volume 4273 of *Lecture Notes in Computer Science*, pages 501–514. Springer, 2006.
- [206] E. Motta. *Reusable Components for Knowledge Modelling*. FAIA-Series. IOS Pres, Amsterdam NL, 1999.
- [207] E. Motta and Z. Zdrahal. Parametric design problem solving. In *Proceedings of the 10th Banff Knowledge Acquisition for Knowledge-Based System Workshop*, 1996.
- [208] B. Nebel. Syntax-based approaches to belief revision. In P. Gärdenfors, editor, *Belief Revision*, volume 29, pages 52–88. Cambridge University Press, Cambridge, UK, 1992.
- [209] A. Newell. The knowledge level. *Artificial Intelligence*, 18:87–127, 1982.
- [210] N. Noy and M. Musen. Ontology versioning in an ontology management framework. *IEEE Intelligent Systems*, 19(4), August 2004.
- [211] D. Nute. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 353–395. Oxford University Press, 1994.

- [212] K. Opalek and J. Wolenski. Normative systems, permission and deontic logic. *Ratio Juris*, 4:334–348, 1991.
- [213] O. Pacheco and J. Carmo. A role based model for the normative specification of organized collective agency and agents interaction. *Journal of Autonomous Agents and Multi-Agent Systems*, 6:145–184, 2003.
- [214] M. Pagnucco. *The Role of Abductive Reasoning within the Process of Belief Revision*. PhD thesis, University of Sydney, 1996.
- [215] L. Pallipoli, F. Pirri, and C. Pizzuti. Algorithms for Selective Enumeration of Prime Implicants. *Artificial Intelligence*, 111:41–72, 1999.
- [216] M. Palmirani. Time model in normative information system. In *Post-proceedings of the ICAIL Workshop on the Role of Legal Knowledge in e-Government*, 2005.
- [217] M. Palmirani. Dynamics of norms over time: a model for legislative consolidation. In *Proceedings of 3rd Legislative XML Workshop, Furore 6-8 April 2005*, Quaderno 18, pages 42–57. CNIPA, 2006.
- [218] M. Palmirani and R. Brighi. Time model for managing the dynamic of normative system. In M. Wimmer, H. Scholl, A. Grönlund, and K. Andersen, editors, *Proceedings of 5th International Conference (EGOV 2006), Krakow, Poland, September 4-8, 2006*, Lecture Notes in Computer Science, pages 207–218. Springer, Berlin, 2006, 2006.
- [219] M. Palmirani, G. Sartor, R. Rubino, A. Boer, E. de Maat, F. Vitali, and E. Francesconi. Guidelines for applying the new format. Deliverable 3.2, Estrella, 2007.
- [220] P. Patel-Schneider and I. Horrocks. A comparison of two modelling paradigms in the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(4):240–250, December 2007.
- [221] M. Peacock. Everyday conceptions of justice and their role in economics. In *Proceedings of the 11th World Congress of Social Economics*, Albertville, 2004.
- [222] A. Pease and I. Niles. IEEE Standard Upper Ontology: A progress report. *Knowledge Engineering Review*, 17:65–70, 2002. Special Issue on Ontologies and Agents.
- [223] R. Pelánek. *Reduction and Abstraction Techniques for Model Checking*. PhD thesis, Masaryk University, Brno, 2006.
- [224] J. Perry. Symposium papers, comments, and an abstract: Cognitive significance and new theories of reference. *Nous*, 22(1):1–18, mar 1988.
- [225] S. Pinker. *The Stuff of Thought: Language As a Window Into Human Nature*. Penguin Group (Viking Press), 2007.
- [226] H. Pinto, A. Prez, and J. Martins. Some Issues on Ontology Integration. In V. Benjamins, B. Chandrasekaran, A. Gomez-Perez, N. Guarino, and M. Uschold, editors, *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5)*, Stockholm, Sweden, 1999.
- [227] P. Piwek. *Logic, Information, Conversation*. PhD thesis, Eindhoven University of Technology, 1998.
- [228] J. L. Pollock. Defeasible reasoning. *Cognitive Science*, 11:481–518, 1987.
- [229] H. Prakken, C. Reed, and D. Walton. Dialogues about the burden of proof. In *ICAIL '05: Proceedings of the 10th international conference on Artificial intelligence and law*, pages 115–124, New York, NY, USA, 2005. ACM.

- [230] H. Prakken and G. Sartor. On the relation between legal language and legal argument: assumptions, applicability and dynamic priorities. In *ICAAIL '95: Proceedings of the 5th international conference on Artificial intelligence and law*, pages 1–10, New York, NY, USA, 1995. ACM.
- [231] H. Prakken and J. Schrickx. Isomorphic models for rules and exceptions in legislation. In J. Breuker, R. de Mulder, and J. Hage, editors, *Legal Knowledge Based Systems. Model-Based Legal Reasoning (Proceedings of JURIX91)*, pages 17–27, Lelystad, the Netherlands, 1991. Koninklijke Vermande.
- [232] H. Prakken and M. Sergot. Dyadic deontic logic and contrary-to-duty obligations. In D. N. Nute, editor, *Defeasible Deontic Logic*, Synthese Library, pages 223–263. Kluwer, 1997.
- [233] A. Rector, N. Drummond, M. Horridge, J. Rogers, H. Knublauch, R. Stevens, H. Wang, and C. Wroe. OWL pizzas: Practical experience of teaching OWL-DL: Common errors & common patterns. In *Proceedings of the European Conference on Knowledge Acquisition (LNAI 3257)*, pages 63–81, Northampton, England, 2004. Springer-Verlag.
- [234] N. Rescher. Axioms for the part relation. *Philosophical Studies*, 6:8–11, 1955.
- [235] E. Rissland and K. Ashley. A note on dimensions and factors. *Artificial Intelligence and Law*, 10:65–77, 2004.
- [236] E. Rissland, K. Ashley, and R. Loui. Ai and law: A fruitful synergy. *Artificial Intelligence*, 150:1–15, 2003.
- [237] E. Rissland and T. Friedman. Detecting change in legal concepts. In *Proceedings of the Fifth International Conference on Artificial Intelligence and Law (ICAAIL-99)*, pages 127–136, New York (NY), 1995. ACM.
- [238] T. Routen. Hierarchically organised formalisations. In *ICAAIL '89: Proceedings of the 2nd international conference on Artificial intelligence and law*, pages 242–250, New York, NY, USA, 1989. ACM.
- [239] L. Royakkers. Giving permission implies giving choice. In *Proceedings of the 8th International Conference on Database and Expert System Applications (DEXA '97)*, 1997.
- [240] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 2nd edition edition, 2003.
- [241] M. Sagoff. On the relation between preference and choice. *Journal of Socio-Economics*, 31(6):587–598, 2003.
- [242] W. S. Sarle. Measurement theory: Frequently asked questions. See <ftp://ftp.sas.com/pub/neural/measurement.faq>, 1995.
- [243] G. Sartor. Why agents comply with norms, and why they should. In R. Conte and C. Dellarocas, editors, *Social Order in MAS*, Dordrecht, 2002. Kluwer.
- [244] G. Sartor. Fundamental legal concepts: A formal and teleological characterisation. Technical report, European University Institute, Florence / Cirsfid, University of Bologna, 2006.
- [245] K. G. Saur. Functional requirements for bibliographic records. *UBCIM Publications - IFLA Section on Cataloguing*, 19, 1998.
- [246] R. Schank and R. Abelson. *Scripts, Plans Goals and Understanding*. Lawrence Erlbaum, New Jersey, 1977.

- [247] P. Schlag. Hiding the ball. *New York University Law Review*, 1681, 1996.
- [248] J. Searle and D. Vanderveken. *Foundations of illocutionary logic*. Cambridge University Press, Cambridge, England, 1985.
- [249] J. R. Searle. *Speech Acts*. Cambridge University Press, 1970.
- [250] M. Sergot and F. Richards. On the representation of action and agency in the theory of normative positions. *Fundamenta Informaticae*, 48:273–293, 2001.
- [251] J. C. Shepherdson. Negation as failure. ii. *J. Log. Program.*, 2(3):185–202, 1985.
- [252] H. Simons. *Personal Income Taxation*. University of Chicago Press, Chicago, IL, USA, 1938.
- [253] E. Sirin, B. C. Grau, and B. Parsia. From wine to water: Optimizing description logic reasoning for nominals. In *KR*, pages 90–99, 2006.
- [254] B. Smith. An essay in formal ontology. *Grazer Philosophische Studien*, 6:39–62, 1978.
- [255] B. Smith and C. Welty. Ontology: Towards a new synthesis. In *Proceedings of Formal Ontology and Information Systems (FOIS)*. Acm, 2001.
- [256] J. Sombekke, T. van Engers, and H. Prakken. Argumentation structures in legal dossiers. In R. Winkels, editor, *Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL-2007)*, New York, 2007. ACM Press.
- [257] M. V. Someren, Y. Barnard, and J. Sandberg. *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*. Academic Press, 1994.
- [258] L. Steels. Components of expertise. *AI Mag.*, 11(2):30–49, 1990.
- [259] J. Straach and K. Truemper. Learning to Ask Relevant Questions. *Artificial Intelligence*, 111:301–328, 1999.
- [260] P. Suber. *The Paradox of Self-Amendment: A Study of Law, Logic, Omnipotence, and Change*. Peter lang Publishing, New York, 1990.
- [261] L. A. Suchman. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge, UK, 1987.
- [262] Y. Tan and L. W. N. van der Torre. How to combine ordering and minimizing in a deontic logic based on preferences. In M. A. Brown and J. Carmo, editors, *Deontic Logic, Agency and Normative Systems: ΔEON '96, Third International Workshop on Deontic Logic in Computer Science, Sesimbra, Portugal, 11–13 January 1996*, pages 216–232. Springer-Verlag, Berlin, 1996.
- [263] F. Tanghe. *Links is soms rechts: over gelijkheid, democratie, en multiculturalisme*. Houtekiet, 2004.
- [264] A. Teije, F. van Harmelen, A. Th, and S. Wielinga. Construction of problem solving methods as parametric design. *International Journal of Human-Computer Studies, special issue on problem-solving methods*, 1998.
- [265] A. Valente. *Legal knowledge engineering: A modelling approach*. PhD thesis, University of Amsterdam, 1995.
- [266] A. Valente and J. Breuker. ON-LINE: An architecture for modelling legal information. In *International Conference on Artificial Intelligence and Law (ICAIL-1995)*, pages 307–315, 1995.
- [267] J. van Benthem, S. van Otterloo, and O. Roy. Preference logic, conditionals, and solution concepts in games. Technical report pp-2005-28, Institute for Logic, Language, and Computation, University of Amsterdam, 2005.

- [268] L. van der Torre. *Reasoning about Obligations: Defeasibility in Preference-based Deontic Logic*. PhD thesis, Erasmus University, Rotterdam, 1997.
- [269] L. W. N. van der Torre and Y. Tan. The temporal analysis of Chisholm's paradox. In T. Senator and B. Buchanan, editors, *Proceedings of the Fourteenth National Conference on Artificial Intelligence and the Ninth Innovative Applications of Artificial Intelligence Conference*, pages 650–655, Menlo Park, California, 1998. AAAI Press.
- [270] T. van Engers, R. Gerrits, M. Boekenoogen, E. Glassée, and P. Kordelaar. POWER: Using UML/OCL for modeling legislation - an application report. In *Proceedings of the International Conference on Artificial Intelligence and Law*. ACM Press, 2001. 1-58113-368-5/01/0005.
- [271] T. van Engers, R. Winkels, A. Boer, and E. de Maat. Knowledge management and the dutch legal aid service counter. In J. J. Schreinemakers and T. van Engers, editors, *Advances in Knowledge Management*, volume IV, Würzburg, 2006. Ergon Verlag.
- [272] T. M. van Engers, R. Winkels, A. Boer, and E. de Maat. Internet, portal to justice? In T. Gordon, editor, *Legal Knowledge and Information Systems. Jurix 2004: The Seventeenth Annual Conference*, Frontiers in Artificial Intelligence and Applications, pages 131–140, Amsterdam, 2004. IOS Press.
- [273] C. van Laer. The Applicability of Comparative Concepts. *European Journal of Comparative Law*, 2.2, 1998.
- [274] J. Velcin and J.-G. Ganascia. Stereo type extraction with default clustering. In *Proc. of IJCAI 2005*, pages 883–888, 2005.
- [275] P. R. S. Visser. *Knowledge Specification for Multiple Legal Tasks. A Case Study of the Interaction Problem in the Legal Domain*. PhD thesis, Leiden University, Leiden, 1995.
- [276] F. Vitali, A. D. Iorio, , and D. Gubellini. Design patterns for document substructures. In *Extreme Markup 2005 Conference. Montreal, 1-5 August 2005*, 2005.
- [277] W. A. Wagenaar. The proper seat: A bayesian discussion of the position of expert witnesses. *Law and Human Behavior*, 12(4):499–510, 1988.
- [278] D. Walton. *Argumentation Methods for Artificial Intelligence in Law*. Springer, 2005.
- [279] D. M. West. Global e-government, 2007 report. Technical report, 2007. <http://www.insidepolitics.org/egovt07int.pdf>.
- [280] B. Wielinga, J. Akkermans, and A. Schreiber. A formal analysis of parametric design problem solving. In *In Proceedings of the 9th Banff Knowledge Acquisition Workshop (KAW95)*, pages 1–15, 1995.
- [281] B. J. Wielinga and G. Schreiber. Configuration-design problem solving. *IEEE Expert*, 12(2):49–56, 1997.
- [282] R. Winkels, A. Boer, E. de Maat, T. van Engers, M. Breebaart, and H. Melger. Constructing a semantic network for legal content. In A. Gardner, editor, *Proceedings of the Tenth International Conference on Artificial Intelligence and Law (ICAAIL)*, pages 125–140, Bologna, Italy, June 2005. IAAAIL, ACM Press.
- [283] R. Winkels, A. Boer, E. de Maat, T. van Engers, M. Breebaart, and H. Melger.

- Constructing a semantic network for legal content. In K. Verbeeck, K. Tuyls, A. Nowé, B. Manderick, and B. Kuipers, editors, *Proceedings of the Seventeenth Belgium-Netherlands Conference on Artificial Intelligence*, pages 405–406. KVAK, 2005. D/2005/0455/15.
- [284] R. Winkels, A. Boer, and R. Hoekstra. CLIME: Lessons learned in Legal Information Serving. In F. van Harmelen, editor, *Proceedings of the European Conference on Artificial Intelligence 2002, Lyon*, Amsterdam, 2002. IOS-Press.
- [285] R. Winkels, A. Boer, and R. Hoekstra. *METAlex*: An XML Standard for Legal Documents. In *Proceedings of the XML Europe Conference*, London (UK), 2003.
- [286] R. Winkels, D. Bosscher, A. Boer, and J. Breuker. Generating exception structures for legal information serving. In T. Gordon, editor, *Proceedings of the Seventh International Conference on Artificial Intelligence and Law (ICAIL-99)*, pages 182–195, New York (NY), 1999. ACM.
- [287] R. Winkels and H. de Bruijn. Making a case for case frames. *Information & Communications Technology Law*, 7(2):117–133, 1998.
- [288] L. Wittgenstein. *Philosophical Investigations*. Blackwell Publishers, April 1973.
- [289] S. E. Zijlstra. *Mandaat en delegatie*. Boom, 2001.

Index

- abducible, 43, 50
- abduction, 42
- ability, 134, 137
- abox, 67
- abstraction, 37, 54, 203, 257
- action, 41, 83, 96, 99, 128
- actor, 244
- adequate, 19
- ADL, 246
- administrative burden, 15, 60
- adversarial, 17, 50
- affirmation, 251
- affirmation of a permission, 208
- affirmation of an imperative, 208, 242
- agency, 123
- agent, 114, 134, 136, 137, 219, 242, 244
- agent causation, 125
- agent role, 271
- Akoma Ntoso, 166
- allowed, 117, 121, 209, 233, 256
- also-applicability, 161, 193
- ampliative reasoning, 37, 42, 58, 74
- applicability, 160
- applicability rule, 252, 268
- argument, 66, 75
- argumentation, 48
- ascription, 46
- assertion, 67
- assumption, 27, 28, 43
 - closed world, 28
 - closure, 72
 - compliance is normal, 28
 - deadline, 84
 - disjointness of universals and particulars, 69, 77
 - frame, 28, 79, 83
 - knowledge availability, 42, 49, 79
 - open world, 64, 70, 72, 108
 - rationality, 25
 - shared mental world, 36
 - single fault, 30
 - unique names, 70
- assumption of closure, 72
- attempt, 127
- autoepistemic, 42
- auxiliary fact, 164
- auxiliary proposition, 149
- auxiliary rule, 163
- axis of reference, 198
- backing, 80, 150
- belief, 47
- belief base, 65, 77, 84
- belief base revision, 66
- belief revision, 65, 228
- belief set, 65
- best explanation, 42
- better, 118
- betterness, 83, 219, 240
- bibliographic citation, 173
- bibliographic expression, 173
- bibliographic item, 173
- bibliographic manifestation, 173
- bibliographic object, 173
- bibliographic work, 173
- bisimulation, 37, 241
- bring about, 129
- brute fact, 93
- brute reality, 93, 106, 251
- burden of proof, 108, 156, 158, 212
- business process, 246
- case, 29
- case description, 248

- case law, 251
- ceteris paribus preference, 221
- choice, 45
- choice rule, 164, 258, 268
- citation, 185
- claim, 48
- closed world assumption, 28, 73
- cognition, 258
- coherent, 34
- collective, 92
- collective entity, 142
- combinative preference, 242
- comparison theory, 198
- competence, 101, 134
- complete, 19
- compliance conflict, 201, 205, 207, 210
- compliance dilemma, 238
- computational legal theory, 35
- concept, 33, 68
- conflict, 211
- conflict of disaffirmation, 205, 220
- conflict with legislator's intentions, 209
- constitutive act, 95, 237
- constitutive rule, 89, 93
- constraint, 73, 75
- context, 34
- contrary-to-duty, 212
- contrary-to-duty obligation, 118
- contrast, 198
- cooperativeness, 17
- core ontologies, 38
- core ontology, 79
- core theories, 38
- CTD, 212
- CTD-obligation, 212

- datalog, 73
- deadline, 84, 130, 243
- deadweight loss, 15
- decidable, 19
- deep model, 280
- default, 66, 75
- default reasoning, 75
- default rule, 75, 245
- defeasibility, 43
- defeasible, 8, 51
- defeasible extension to a knowledge component, 67
- definition, 33
- delayed application, 153
- delegation, 142, 149, 161, 164, 189, 192, 206
- denotational semantics, 20
- deontic choice, 44, 200
- deontic logic, 26, 116, 118
- deontic square, 137
- description classifier, 72
- description logic, 8, 67
- descriptive fidelity, 167, 191
- design, 84
- design principle, 19
- disability, 138, 141
- disaffirmation of a permission, 205
- disaffirmation of an imperative, 205
- disallowed, 117, 122, 209, 233, 256
- disjointness, 39
- disjointness of universals and particulars
 - assumption, 69
- dispositional beliefs, 65
- domain, 22
- double jeopardy, 231
- doxastic, 25, 46, 50
- doxastic voluntarism, 46, 60
- duty, 139
- duty conditional on violation, 216
- dynamic modal logic, 84

- eGovernment, 15, 16
- emergence, 128
- endophoric proform expression, 184, 188
- envisioning, 256
- epistemic, 32
- epistemic competence, 28, 191
- epistemic entrenchment, 44, 51, 228
- epistemic obligation, 231, 235
- epistemic process, 41
- epistemic role, 30, 47, 109
- epistemic thing, 21
- epistemics, 31
- epistemological, 32

- epistemological gap, 29, 81
- epistemology, 7, 31, 40
- equality of conditions, 114
- evaluative, 116, 250
- evaluative concept, 255
- event, 83
- event calculus, 83, 128
- everyone, 142
- ex ante, 47, 83, 255
- ex post, 47, 83, 255
- ex post facto, 153
- ex tunc, 154, 169, 180, 193
- exclusionary preference, 242
- exclusive jurisdiction, 165
- exophoric proform expression, 184
- expectation, 109, 110, 114, 132, 136
- expected utility, 45
- expected utility hypothesis, 45, 242
- explanation, 43, 48, 75
- explanatory power, 43
- explicit disaffirmation, 207
- explicit knowledge, 16
- expressiveness, 19
- extension, 20
- externalism, 25, 46

- factor, 199, 249, 250, 257
- falsifiability, 26
- feature engineering, 24
- fidelity, 20, 52
- first representation, 189, 266
- fluent, 83, 128
- folk psychology, 25, 82
- formal act, 102
- frame assumption, 28, 83
- FRBR, 82, 190
- function, 114
- Functional Ontology of Law, 134
- Functional Requirements for Bibliographic Records, 82

- generic case, 210, 250
- granularity, 36, 83
- ground facts, 73
- Grundnorm, 90

- hierarchical task network, 98
- Hoare logic, 84
- Hohfeld, 137, 243
- homonymy, 35
- horizontal equity, 255
- human-based computation, 49
- hybrid, 72
- hypothesis, 21, 42, 43
- hypothetical object, 78

- I-intention, 95, 245
- I-preference, 245
- imitation, 245
- immunity, 141
- imperium, 164
- impossible to realize, 209
- impotence, 138, 141
- incision function, 66, 78, 233, 235
- inclusion, 186
- incompetence, 138, 141
- indicative rule, 156
- individual, 67
- individuals, 34
- inference, 23, 42
- information, 34, 49
- initiation, 83
- institution, 89, 92, 265
- institutional reality, 90, 93, 106
- institutional rule, 93, 267
- intended normative order, 88
- intension, 21
- intensional semantics, 21
- intent, 156
- intention, 92, 94–96, 124, 130, 137, 153, 162, 197, 259
- intention entrenchment, 44
- intentional stance, 21
- intentional thing, 21, 82
- intercoder reliability, 191
- interest, 136
- internalism, 46
- intersection conflict, 206, 234, 235, 259
- introspection, 21, 76
- isomorphism, 7, 19, 20, 29, 54, 81, 185, 188

- item, 151, 178
- jointly realizable, 205
- judicial review, 193
- jural opposite, 137
- jural relationship, 137, 139
- jurisdiction, 164
- jurisprudence, 248
- jurisprudence constante, 248
- justification, 48
- KBS, 14, 23, 34, 36, 41, 49, 73, 78, 108
- know-how, 40
- know-why, 40
- knowledge acquisition, 198
- knowledge as a productive factor, 29, 30
- knowledge as the attainment of truth, 29
- knowledge base, 17, 18, 67
- knowledge based systems, 14
- knowledge component, 28, 65
- knowledge domain, 32
- knowledge level hypothesis, 25
- knowledge representation, 17
- knowledge representation language, 18
- knowledge source, 64, 65
- knowledge syndication, 29, 61
- legal act, 100, 232
- legal fact, 100, 105, 188
- legal fiction, 161, 193
- legal knowledge based system, 15
- legal knowledge engineering, 16
- legal ontology, 36, 100, 254
- legal planning, 279
- legal pluralism, 4
- legal rule, 100, 188, 248, 266
- legal writing, 80
- legally relevant, 148
- legislation, 88
- legislative act, 102
- legislative drafting, 125
- legislative resource, 80
- legislative supremacy, 193
- legislator, 101, 113, 122, 125, 163
- legitimate expectation, 136
- lex posterior, 149, 189, 229
- lex specialis, 149, 229, 233, 251
- lex superior, 149, 229
- liability, 141
- LKBS, 15, 29, 78, 156, 158, 188, 191, 199, 237, 258
- LKIF, 79, 150, 200, 217, 235, 282
- logic, 19
- logical nonmonotonicity, 229
- man on the Clapham Omnibus, 127
- mandate, 164
- meaning, 22
- measurement theory, 198
- mens rea, 47
- mental object, 247
- mental things, 20, 25
- merging concepts, 192
- metadata, 61, 149, 151, 167
- metalegal knowledge, 135
- MetaLex, 82, 149, 150, 152, 166, 173, 248, 252, 264, 282
- metaphor, 35, 36
- methodological priority of ex ante intention over ex post intention, 96
- metonymy, 35
- mimetic theory, 245
- minimal explanation, 43
- minimal model, 75
- minimality, 59
- model, 72
- monotonic, 19
- monotonicity, 19, 64
- motivating indifference, 241
- motivating preference, 219, 241
- naive physics, 82
- namespace, 62, 67
- ne bis in idem, 208, 231, 254
- negation as failure, 73
- negative causation, 128, 132
- no-dilemma assumption, 222
- nominal, 68
- nonmonotonicity, 19
- norm, 26, 91, 109, 218, 240
- norm of analysis, 113, 242, 249, 255
- norm qua qualification, 203

- normality, 100, 109
- normative conditional, 128
- normative conflict, 203, 258
- normative order, 54, 109
- normative position, 133
- normative proposition, 26
- normative rule, 89, 112, 116–118, 162, 200, 258, 268, 269
- normative statement, 26
- normativity, 109
- Norme in Rete, 166
- nothing, 21
- nulla poena sine lege, 153

- obligation, 90, 115, 122, 124, 129, 140
- occurrent beliefs, 65
- ontic, 32
- ontological, 86
- ontological commitment, 250
- ontological commitments, 22
- ontological distinction principle, 39
- ontological stratification, 36, 38, 53, 86
- ontology, 7, 23, 32
- ontology integration, 254
- open texture, 22
- operational semantics, 20
- opinio iuris, 251
- opposite, 210
- ought-to-be, 124
- ought-to-do, 124
- OWL, 4, 7, 64, 67
- OWL DL, 7, 20, 67

- parliamentary sovereignty, 193
- participant, 168
- patient, 244
- PDDL, 246
- perfect fidelity, 21
- perfect refinement, 37
- period, 129
- permission, 90, 115, 123, 140
- personal construct, 198
- physical things, 20
- plan, 96
- plan execution, 97, 277
- plan recognition, 88

- planning, 88, 98
- polysemy, 35
- postdiction, 47
- potestative right, 101, 137, 140, 142
- potestative square, 138
- power, 101, 133, 134, 141
- prediction, 47
- preference, 44, 45, 83, 91, 97, 110, 115, 119, 136, 137, 218
- preference independence, 242
- preference logic, 119
- preference theory, 218, 241
- prime implicant, 66, 208, 210
- primum comparandum, 198
- principle of rationality, 25
- privilege, 140
- problem solving, 14, 40, 52
- problem solving competence, 23, 28, 52
- problem solving method, 28, 42, 155
- problem type, 28
- prohibition, 90, 115, 122
- proof tree, 50
- proposition, 65
- propositional attitude, 65
- prospect theory, 110
- prototype, 37, 246
- psychological geometry, 198
- punishment, 116

- query language, 72

- rational choice theory, 46
- RDF, 62
- reaction, 134
- reactive knowledge, 134
- realizability, 205
- reasoning as process, 84
- reduction, 37
- reductionism, 36
- reference, 266
- reference transaction, 110, 111
- reflexivity, 219
- regime, 253
- reification, 47
- repertory grid, 120
- representation, 266

- requirement, 252
- resource, 63
- restorative justice, 135
- retroactive applicability, 153
- reusability, 29
- revealed preference, 221, 240
- rhetoric, 51
- rhetorical comparison, 198
- rhetorical hierarchy, 51
- right, 139–141
- role, 91, 125
- rule, 248

- scalability, 27
- scale, 198
- schema, 33
- script, 100, 111, 139, 243, 245
- secundum comparatum, 198
- self, 246
- Semantic Web, 15, 29, 60
- settled context, 207, 212, 224
- settled model, 78
- shared jurisdiction, 165
- sichttag, 154, 169
- silent, 117, 209
- simile, 198
- simulation, 37
- simulation theory approach to the mind, 97
- single fault assumption, 30
- situated action, 29
- situation, 128, 131
- skeletal plan, 112
- skill, 29
- social, 92, 265
- social role, 113
- soundness, 19
- source of law, 16, 80, 150, 264
- sources of law, 27
- spatial planning, 82
- speech act, 47
- stance, 36
- stare decisis, 14, 231, 247
- stare dictis, 248
- stare rationibus decidendi, 248

- state, 128
- state of obligation, 117, 133
- stative, 134
- status quo bias, 110
- stichtag, 154, 169
- stratum, 38
- strengthening the antecedent, 212, 226
- STRIPS, 246
- strong permission, 201, 209
- subdelegation, 164
- subjective norm, 246
- subjunctive betterness, 118, 119, 198, 255, 270
- submandate, 164
- surrogate, 20
- synecdoche, 35

- tacit knowledge, 16, 29, 81
- TBox, 219
- tbox, 67
- temporal nonmonotonicity, 229
- term, 33, 188, 266
- termination, 83
- terminological subsumption, 34, 37
- terminology, 33
- tertium comparationis, 198
- theft, 116
- theory of brute reality, 211
- theory-theory approach to the mind, 97
- top ontology, 82
- tractable, 19
- trichotomy, 221

- uniform resource identifier, 7, 61
- uniform resource locator, 174
- unintended but goal-directed behaviour, 99
- unique bibliographic identifier, 173
- unique name assumption, 58, 73
- unique names assumption, 70
- URI, 7, 61
- URI reference, 62
- utility theory, 45, 110

- validity, 19
- victim, 135

violation, 115, 122, 133

we-intention, 95, 245

we-preference, 245

weak permission, 201, 209

weakening the consequent, 226

Web Ontology Language, 64

xml base, 178, 190

zeugma, 35

SIKS Dissertation Series

- 1998-1** Johan van den Akker (CWI). DE GAS - An Active, Temporal Database of Autonomous Objects.
- 1998-2** Floris Wiesman (UM). Information Retrieval by Graphically Browsing Meta-Information.
- 1998-3** Ans Steuten (TUD). A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective.
- 1998-4** Dennis Breuker (UM). Memory versus Search in Games.
- 1998-5** E.W.Oskamp (RUL). Computerondersteuning bij Straftoemeting.
- 1999-1** Mark Sloof (VU). Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products.
- 1999-2** Rob Potharst (EUR). Classification using decision trees and neural nets.
- 1999-3** Don Beal (UM). The Nature of Minimax Search.
- 1999-4** Jacques Penders (UM). The practical Art of Moving Physical Objects.
- 1999-5** Aldo de Moor (KUB). Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems.
- 1999-6** Niek J.E. Wijngaards (VU). Re-design of compositional systems.
- 1999-7** David Spelt (UT). Verification support for object database design.
- 1999-8** Jacques H.J. Lenting (UM). Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation.
- 2000-1** Frank Niessink (VU). Perspectives on Improving Software Maintenance.
- 2000-2** Koen Holtman (TUE). Prototyping of CMS Storage Management.
- 2000-3** Carolien M.T. Metselaar (UVA). Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief.
- 2000-4** Geert de Haan (VU). ETAG, A Formal Model of Competence Knowledge for User Interface Design.
- 2000-5** Ruud van der Pol (UM). Knowledge-based Query Formulation in Information Retrieval.
- 2000-6** Rogier van Eijk (UU). Programming Languages for Agent Communication.
- 2000-7** Niels Peek (UU). Decision-theoretic Planning of Clinical Patient Management.
- 2000-8** Veerle Coup (EUR). Sensitivity Analysis of Decision-Theoretic Networks.
- 2000-9** Florian Waas (CWI). Principles of Probabilistic Query Optimization.
- 2000-10** Niels Nes (CWI). Image Database Management System Design Considerations, Algorithms and Architecture.
- 2000-11** Jonas Karlsson (CWI). Scalable Distributed Data Structures for Database Management.
- 2001-1** Silja Renooij (UU). Qualitative Approaches to Quantifying Probabilistic Networks.
- 2001-2** Koen Hindriks (UU). Agent Programming Languages: Programming with Mental Models.
- 2001-3** Maarten van Someren (UvA). Learning as problem solving.
- 2001-4** Evgueni Smirnov (UM). Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets.
- 2001-5** Jacco van Ossenbruggen (VU). Processing Structured Hypermedia: A Matter of Style.
- 2001-6** Martijn van Welie (VU). Task-based User Interface Design.
- 2001-7** Bastiaan Schonhage (VU). Diva: Architectural Perspectives on Information Visualization.
- 2001-8** Pascal van Eck (VU). A Compositional Semantic Structure for Multi-Agent Systems Dynamics.
- 2001-9** Pieter Jan 't Hoen (RUL). Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes.
- 2001-10** Maarten Sierhuis (UvA). Modeling and Simulating Work Practice. BRAHMS: a multiagent modeling and simulation language for work practice analysis and design.
- 2001-11** Tom M. van Engers (VUA). Knowledge Management: The Role of Mental Models in Business Systems Design.
- 2002-01** Nico Lassing (VU). Architecture-Level Modifiability Analysis.
- 2002-02** Roelof van Zwol (UT). Modelling and searching web-based document collections.
- 2002-03** Henk Ernst Blok (UT). Database Optimization Aspects for Information Retrieval.

- 2002-04** Juan Roberto Castelo Valdueza (UU). The Discrete Acyclic Digraph Markov Model in Data Mining.
- 2002-05** Radu Serban (VU). The Private Cyberspace Modeling Electronic Environments inhabited by Privacy-concerned Agents.
- 2002-06** Laurens Mommers (UL). Applied legal epistemology; Building a knowledge-based ontology of the legal domain.
- 2002-07** Peter Boncz (CWI). Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications.
- 2002-08** Jaap Gordijn (VU). Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas.
- 2002-09** Willem-Jan van den Heuvel(KUB). Integrating Modern Business Applications with Objectified Legacy Systems.
- 2002-10** Brian Sheppard (UM). Towards Perfect Play of Scrabble.
- 2002-11** Wouter C.A. Wijngaards (VU). Agent Based Modelling of Dynamics: Biological and Organisational Applications.
- 2002-12** Albrecht Schmidt (UVA). Processing XML in Database Systems.
- 2002-13** Hongjing Wu (TUE). A Reference Architecture for Adaptive Hypermedia Applications.
- 2002-14** Wieke de Vries (UU). Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems.
- 2002-15** Rik Eshuis (UT). Semantics and Verification of UML Activity Diagrams for Workflow Modelling.
- 2002-16** Pieter van Langen (VU). The Anatomy of Design: Foundations, Models and Applications.
- 2002-17** Stefan Manegold (UVA). Understanding, Modeling, and Improving Main-Memory Database Performance.
- 2003-01** Heiner Stuckenschmidt (VU). Ontology-Based Information Sharing in Weakly Structured Environments.
- 2003-02** Jan Broersen (VU). Modal Action Logics for Reasoning About Reactive Systems.
- 2003-03** Martijn Schuemie (TUD). Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy.
- 2003-04** Milan Petkovic (UT). Content-Based Video Retrieval Supported by Database Technology.
- 2003-05** Jos Lehmann (UVA). Causation in Artificial Intelligence and Law - A modelling approach.
- 2003-06** Boris van Schooten (UT). Development and specification of virtual environments.
- 2003-07** Machiel Jansen (UvA). Formal Explorations of Knowledge Intensive Tasks.
- 2003-08** Yongping Ran (UM). Repair Based Scheduling.
- 2003-09** Rens Kortmann (UM). The resolution of visually guided behaviour.
- 2003-10** Andreas Lincke (UvT). Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture.
- 2003-11** Simon Keizer (UT). Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks.
- 2003-12** Roeland Ordelman (UT). Dutch speech recognition in multimedia information retrieval.
- 2003-13** Jeroen Donkers (UM). Nosce Hostem - Searching with Opponent Models.
- 2003-14** Stijn Hoppenbrouwers (KUN). Freezing Language: Conceptualisation Processes across ICT-Supported Organisations.
- 2003-15** Mathijs de Weerd (TUD). Plan Merging in Multi-Agent Systems
- 2003-16** Menzo Windhouwer (CWI). Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouses.
- 2003-17** David Jansen (UT). Extensions of Statecharts with Probability, Time, and Stochastic Timing.
- 2003-18** Levente Kocsis (UM). Learning Search Decisions.
- 2004-01** Virginia Dignum (UU). A Model for Organizational Interaction: Based on Agents, Founded in Logic.
- 2004-02** Lai Xu (UvT). Monitoring Multi-party Contracts for E-business.
- 2004-03** Perry Groot (VU). A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving.

- 2004-04** Chris van Aart (UVA). Organizational Principles for Multi-Agent Architectures.
- 2004-05** Viara Popova (EUR). Knowledge discovery and monotonicity.
- 2004-06** Bart-Jan Hommes (TUD). The Evaluation of Business Process Modeling Techniques.
- 2004-07** Elise Boltjes (UM). Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes.
- 2004-08** Joop Verbeek(UM). Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politieke gegevensuitwisseling en digitale expertise.
- 2004-09** Martin Caminada (VU). For the Sake of the Argument; explorations into argument-based reasoning.
- 2004-10** Suzanne Kabel (UVA). Knowledge-rich indexing of learning-objects.
- 2004-11** Michel Klein (VU). Change Management for Distributed Ontologies.
- 2004-12** The Duy Bui (UT). Creating emotions and facial expressions for embodied agents.
- 2004-13** Wojciech Jamroga (UT). Using Multiple Models of Reality: On Agents who Know how to Play.
- 2004-14** Paul Harrenstein (UU). Logic in Conflict. Logical Explorations in Strategic Equilibrium.
- 2004-15** Arno Knobbe (UU). Multi-Relational Data Mining.
- 2004-16** Federico Divina (VU). Hybrid Genetic Relational Search for Inductive Learning.
- 2004-17** Mark Winands (UM). Informed Search in Complex Games.
- 2004-18** Vania Bessa Machado (UvA). Supporting the Construction of Qualitative Knowledge Models.
- 2004-19** Thijs Westerveld (UT). Using generative probabilistic models for multimedia retrieval.
- 2004-20** Madelon Evers (Nyenrode). Learning from Design: facilitating multidisciplinary design teams.
- 2005-01** Floor Verdenius (UVA). Methodological Aspects of Designing Induction-Based Applications.
- 2005-02** Erik van der Werf (UM). AI techniques for the game of Go.
- 2005-03** Franc Grootjen (RUN). A Pragmatic Approach to the Conceptualisation of Language.
- 2005-04** Nirvana Meratnia (UT). Towards Database Support for Moving Object data.
- 2005-05** Gabriel Infante-Lopez (UVA). Two-Level Probabilistic Grammars for Natural Language Parsing.
- 2005-06** Pieter Spronck (UM). Adaptive Game AI.
- 2005-07** Flavius Frasincar (TUE). Hypermedia Presentation Generation for Semantic Web Information Systems.
- 2005-08** Richard Vdovjak (TUE). A Model-driven Approach for Building Distributed Ontology-based Web Applications.
- 2005-09** Jeen Broekstra (VU). Storage, Querying and Inferencing for Semantic Web Languages.
- 2005-10** Anders Bouwer (UVA). Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments.
- 2005-11** Elth Ogston (VU). Agent Based Matchmaking and Clustering - A Decentralized Approach to Search.
- 2005-12** Csaba Boer (EUR). Distributed Simulation in Industry.
- 2005-13** Fred Hamburg (UL). Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen.
- 2005-14** Borys Omelayenko (VU). Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics.
- 2005-15** Tibor Bosse (VU). Analysis of the Dynamics of Cognitive Processes.
- 2005-16** Joris Graaumans (UU). Usability of XML Query Languages.
- 2005-17** Boris Shishkov (TUD). Software Specification Based on Re-usable Business Components.
- 2005-18** Danielle Sent (UU). Test-selection strategies for probabilistic networks.
- 2005-19** Michel van Dartel (UM). Situated Representation.
- 2005-20** Cristina Coteanu (UL). Cyber Consumer Law, State of the Art and Perspectives.
- 2005-21** Wijnand Derks (UT). Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics.
- 2006-01** Samuil Angelov (TUE). Foundations of B2B Electronic Contracting.

- 2006-02** Cristina Chisalita (VU). Contextual issues in the design and use of information technology in organizations.
- 2006-03** Noor Christoph (UVA). The role of metacognitive skills in learning to solve problems.
- 2006-04** Marta Sabou (VU). Building Web Service Ontologies.
- 2006-05** Cees Pierik (UU). Validation Techniques for Object-Oriented Proof Outlines.
- 2006-06** Ziv Baida (VU). Software-aided Service Bundling - Intelligent Methods & Tools for Graphical Service Modeling.
- 2006-07** Marko Smiljanic (UT). XML schema matching – balancing efficiency and effectiveness by means of clustering.
- 2006-08** Eelco Herder (UT). Forward, Back and Home Again - Analyzing User Behavior on the Web.
- 2006-09** Mohamed Wahdan (UM). Automatic Formulation of the Auditor's Opinion.
- 2006-10** Ronny Siebes (VU). Semantic Routing in Peer-to-Peer Systems.
- 2006-11** Joeri van Ruth (UT). Flattening Queries over Nested Data Types.
- 2006-12** Bert Bongers (VU). Interactivation - Towards an e-cology of people, our technological environment, and the arts.
- 2006-13** Henk-Jan Lebbink (UU). Dialogue and Decision Games for Information Exchanging Agents.
- 2006-14** Johan Hoorn (VU). Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change.
- 2006-15** Rainer Malik (UU). CONAN: Text Mining in the Biomedical Domain.
- 2006-16** Carsten Riggelsen (UU). Approximation Methods for Efficient Learning of Bayesian Networks.
- 2006-17** Stacey Nagata (UU). User Assistance for Multitasking with Interruptions on a Mobile Device.
- 2006-18** Valentin Zhizhkun (UVA). Graph transformation for Natural Language Processing.
- 2006-19** Birna van Riemsdijk (UU). Cognitive Agent Programming: A Semantic Approach.
- 2006-20** Marina Velikova (UvT). Monotone models for prediction in data mining.
- 2006-21** Bas van Gils (RUN). Aptness on the Web
- 2006-22** Paul de Vrieze (RUN). Fundaments of Adaptive Personalisation.
- 2006-23** Ion Juvina (UU). Development of Cognitive Model for Navigating on the Web.
- 2006-24** Laura Hollink (VU). Semantic Annotation for Retrieval of Visual Resources.
- 2006-25** Madalina Drugan (UU). Conditional log-likelihood MDL and Evolutionary MCMC.
- 2006-26** Vojkan Mihajlović (UT). Score Region Algebra: A Flexible Framework for Structured Information Retrieval.
- 2006-27** Stefano Bocconi (CWI). Vox Populi: generating video documentaries from semantically annotated media repositories.
- 2006-28** Borkur Sigurbjornsson (UVA). Focused Information Access using XML Element Retrieval.
- 2007-01** Kees Leune (UvT). Access Control and Service-Oriented Architectures.
- 2007-02** Wouter Teepe (RUG). Reconciling Information Exchange and Confidentiality: A Formal Approach.
- 2007-03** Peter Mika (VU). Social Networks and the Semantic Web.
- 2007-04** Jurriaan van Diggelen (UU). Achieving Semantic Interoperability in Multi-agent Systems: a dialogue-based approach.
- 2007-05** Bart Schermer (UL). Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance..
- 2007-06** Gilad Mishne (UVA). Applied Text Analytics for Blogs.
- 2007-07** Natasa Jovanović (UT). To Whom It May Concern - Addressee Identification in Face-to-Face Meetings.
- 2007-08** Mark Hoogendoorn (VU). Modeling of Change in Multi-Agent Organizations.
- 2007-09** David Mobach (VU). Agent-Based Mediated Service Negotiation.
- 2007-10** Huib Aldewereld (UU). Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols.
- 2007-11** Natalia Stash (TUE). Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System.

- 2007-12** Marcel van Gerven (RUN). Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty.
- 2007-13** Rutger Rienks (UT). Meetings in Smart Environments; Implications of Progressing Technology.
- 2007-14** Niek Bergboer (UM). Context-Based Image Analysis.
- 2007-15** Joyca Lacroix (UM). NIM: a Situated Computational Memory Model.
- 2007-16** Davide Grossi (UU). Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems.
- 2007-17** Theodore Charitos (UU). Reasoning with Dynamic Networks in Practice.
- 2007-18** Bart Orriens (UvT). On the development an management of adaptive business collaborations.
- 2007-19** David Levy (UM). Intimate relationships with artificial partners.
- 2007-20** Slinger Jansen (UU). Customer Configuration Updating in a Software Supply Network.
- 2007-21** Karianne Vermaas (UU). Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005.
- 2007-22** Zlatko Zlatev (UT). Goal-oriented design of value and process models from patterns.
- 2007-23** Peter Barna (TUE). Specification of Application Logic in Web Information Systems.
- 2007-24** Georgina Ramrez Camps (CWI). Structural Features in XML Retrieval.
- 2007-25** Joost Schalken (VU). Empirical Investigations in Software Process Improvement.
- 2008-01** Katalin Boer-Sorbn (EUR). Agent-Based Simulation of Financial Markets: A modular, continuous-time approach.
- 2008-02** Alexei Sharpanskykh (VU). On Computer-Aided Methods for Modeling and Analysis of Organizations.
- 2008-03** Vera Hollink (UVA). Optimizing hierarchical menus: a usage-based approach.
- 2008-04** Ander de Keijzer (UT). Management of Uncertain Data - towards unattended integration.
- 2008-05** Bela Mutschler (UT). Modeling and simulating causal dependencies on process-aware information systems from a cost perspective.
- 2008-06** Arjen Hommersom (RUN). On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective.
- 2008-07** Peter van Rosmalen (OU). Supporting the tutor in the design and support of adaptive e-learning.
- 2008-08** Janneke Bolt (UU). Bayesian Networks: Aspects of Approximate Inference.
- 2008-09** Christof van Nimwegen (UU). The paradox of the guided user: assistance can be counter-effective.
- 2008-10** Wauter Bosma (UT). Discourse oriented summarization.
- 2008-11** Vera Kartseva (VU). Designing Controls for Network Organizations: A Value-Based Approach.
- 2008-12** Jozsef Farkas (RUN). A Semiotically Oriented Cognitive Model of Knowledge Representation.
- 2008-13** Caterina Carraciolo (UVA). Topic Driven Access to Scientific Handbooks.
- 2008-14** Arthur van Bunningen (UT). Context-Aware Querying; Better Answers with Less Effort.
- 2008-15** Martijn van Otterlo (UT). The Logic of Adaptive Behavior: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains.
- 2008-16** Henriette van Vugt (VU). Embodied agents from a user's perspective.
- 2008-17** Martin Op 't Land (TUD). Applying Architecture and Ontology to the Splitting and Allying of Enterprises.
- 2008-18** Guido de Croon (UM). Adaptive Active Vision.
- 2008-19** Henning Rode (UT). From Document to Entity Retrieval: Improving Precision and Performance of Focused Text Search.
- 2008-20** Rex Arendsen (UVA). Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven.
- 2008-21** Krisztian Balog (UVA). People Search in the Enterprise.
- 2008-22** Henk Koning (UU). Communication of IT-Architecture.

- 2008-23** Stefan Visscher (UU). Bayesian network models for the management of ventilator-associated pneumonia.
- 2008-24** Zharko Aleksovski (VU). Using background knowledge in ontology matching.
- 2008-25** Geert Jonker (UU). Efficient and Equitable Exchange in Air Traffic Management Plan Repair using Spender-signed Currency.
- 2008-26** Marijn Huijbregts (UT). Segmentation, Diarization and Speech Transcription: Surprise Data Unraveled.
- 2008-27** Hubert Vogten (OU). Design and Implementation Strategies for IMS Learning Design.
- 2008-28** Ildiko Flesch (RUN). On the Use of Independence Relations in Bayesian Networks.
- 2008-29** Dennis Reidsma (UT). Annotations and Subjective Machines - Of Annotators, Embodied Agents, Users, and Other Humans.
- 2008-30** Wouter van Atteveldt (VU). Semantic Network Analysis: Techniques for Extracting, Representing and Querying Media Content.
- 2008-31** Loes Braun (UM). Pro-Active Medical Information Retrieval.
- 2008-32** Trung H. Bui (UT). Toward Affective Dialogue Management using Partially Observable Markov Decision Processes.
- 2008-33** Frank Terpstra (UVA). Scientific Workflow Design; theoretical and practical issues.
- 2008-34** Jeroen de Knijf (UU). Studies in Frequent Tree Mining.
- 2008-35** Ben Torben Nielsen (UvT). Dendritic morphologies: function shapes structure.
- 2009-01** Rasa Jurgelenaite (RUN). Symmetric Causal Independence Models.
- 2009-02** Willem Robert van Hage (VU). Evaluating Ontology-Alignment Techniques.
- 2009-03** Hans Stol (UvT). A Framework for Evidence-based Policy Making Using IT.
- 2009-04** Josephine Nabukenya (RUN). Improving the Quality of Organisational Policy Making using Collaboration Engineering.
- 2009-05** Sietse Overbeek (RUN). Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality.
- 2009-06** Muhammad Subianto (UU). Understanding Classification.
- 2009-07** Ronald Poppe (UT). Discriminative Vision-Based Recovery and Recognition of Human Motion.
- 2009-08** Volker Nannen (VU). Evolutionary Agent-Based Policy Analysis in Dynamic Environments.
- 2009-09** Benjamin Kanagwa (RUN). Design, Discovery and Construction of Service-oriented Systems.
- 2009-10** Jan Wielemaker (UVA). Logic programming for knowledge-intensive interactive applications.
- 2009-11** Alexander Boer (UVA). Legal Theory, Sources of Law & the Semantic Web.
- 2009-12** Peter Massuthe (TUE, Humboldt-Universität zu Berlin). Operating Guidelines for Services.