A rule-based deontic reasoner

X. Parent

RuleML Webinar 26th January 2018
University of Luxembourg

Department of Computer Science

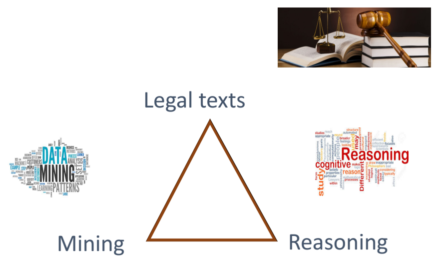
[1] Joint work with L. van der Torre

Talk layout

Background: AI & law

MIREL

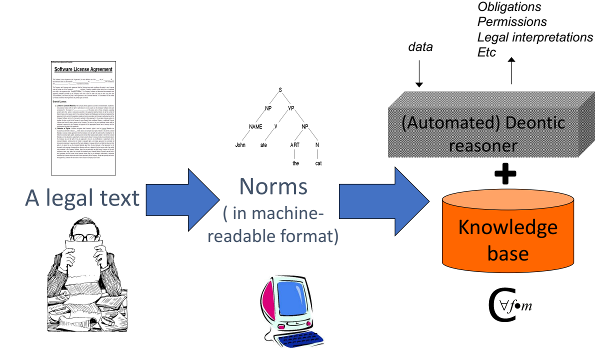
MINING AND REASONING WITH LEGAL TEXTS.



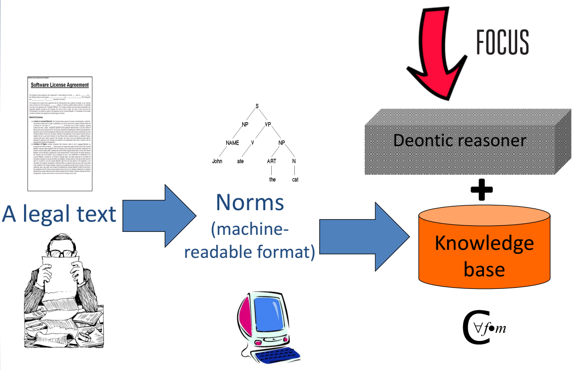
EU Horizon 2020 research and innovation programme–Marie Skodowska-Curie

Grant agreement No 690974.

Mining and reasoning



Mining and reasoning



Long term-goal: automated tool support

- Norm-compliance checking
- Consistency checking

Example


Risk fines of up to
€20 million
or 4% of your organisations
global turnover
How do I Comply with the new EU GDPR?
IG-SMART.COM
General Data Protection Regulation
25th May 2018

Introduction

Example

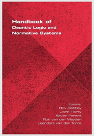
Deontic logic

Deontic logic

- Concerned with obligation, permission and related concepts
- Normative reasoning in law
  - Sergot, McCarthy, Jones, Governatori, Sartor, ...
  - Law as a logical theory

Two research traditions

- Possible worlds semantics (mid 50s)
  - Deontic logic as a branch of modal logic
- “Norm-based” semantics (Hansen, 00s)
  - Roots in Alchourrón and Bulygin’s approach to normative systems



Deontic logic

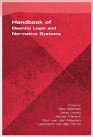
Deontic logic

- Concerned with obligation, permission and related concepts
- Normative reasoning in law
  - Sergot, McCarthy, Jones, Governatori, Sartor, ...
  - Law as a logical theory

Two research traditions

- Possible worlds semantics (mid 50s)
  - Deontic logic as a branch of modal logic
- “Norm-based”semantics (Hansen, 00s)
  - Roots in Alchourrón and Bulygin’s approach to normative systems

Rule-based systems



Introduction

[Parent and van der Torre, 2017]

Parent, X. and van der Torre, L. W. N. (2017).
The pragmatic oddity in a norm-based deontic logic.
In Keppens, J. and Governatori, G., editors, *Proc. of the 16th International Conference on Articial Intelligence and Law, ICAIL 2017, London, United Kingdom, June 12-16, 2017*, pages 169–178.

Highlights

[Parent and van der Torre, 2017]

Parent, X. and van der Torre, L. W. N. (2017).
The pragmatic oddity in a norm-based deontic logic.
In Keppens, J. and Governatori, G., editors, *Proc. of the 16th International Conference on Articial Intelligence and Law, ICAIL 2017, London, United Kingdom, June 12-16, 2017*, pages 169–178.

Highlights

Makinson/van der torre

A "new" logic–in the rule-based tradition (I/O logic)

RuleML 2015

- Well-defined
- Performs well w.r.t. benchmark problems of deontic logic
  - Contrary-to-duty (CTD) reasoning
  - Conflict

(Perform well=return the expected answers to queries)

[Parent and van der Torre, 2017]

Parent, X. and van der Torre, L. W. N. (2017).
The pragmatic oddity in a norm-based deontic logic.
In Keppens, J. and Governatori, G., editors, *Proc. of the 16th International Conference on Articial Intelligence and Law, ICAIL 2017, London, United Kingdom, June 12-16, 2017*, pages 169–178.

Highlights

Advantage

- Simple (on the outside)
- User-friendly

  Easy to use for non-experts

Introduction

[Parent and van der Torre, 2017]

Parent, X. and van der Torre, L. W. N. (2017).
The pragmatic oddity in a norm-based deontic logic.
In Keppens, J. and Governatori, G., editors, *Proc. of the 16th International Conference on Articial Intelligence and Law, ICAIL 2017, London, United Kingdom, June 12-16, 2017*, pages 169–178.

Highlights

Novelty 1

- Consistency checks in the semantics
- Reflected in the proof theory

Spin-off

- Handles a recurrent objection against rule-based systems (e.g., Reiter's default logic): lack of a proof-theory

Introduction

[Parent and van der Torre, 2017]

Parent, X. and van der Torre, L. W. N. (2017).
The pragmatic oddity in a norm-based deontic logic.
In Keppens, J. and Governatori, G., editors, *Proc. of the 16th International Conference on Articial Intelligence and Law, ICAIL 2017, London, United Kingdom, June 12-16, 2017*, pages 169–178.

Highlights

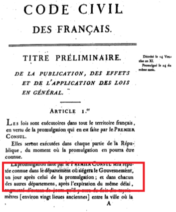
Novelty 2

- A modular treatment of the two categories of benchmarks
  - A unique formalism, not two (separate) formalisms

Spin-off

- The reasoner able to handle **both** CTDs and conflicts in the text

Group 1: CTD



CODE CIVIL
DES FRANÇAIS.

TITRE PRÉLIMINAIRE.

DE LA PUBLICATION, DES EFFETS ET DE L'APPLICATION DES LOIS EN GÉNÉRAL.

Décrété le 14 Ventôse an XI. Promulgué le 24 du même mois.

ARTICLE 1.er

LES lois sont exécutoires dans tout le territoire français, en vertu de la promulgation qui en est faite par le PREMIER CONSUL.

Elles seront exécutées dans chaque partie de la République, du moment où la promulgation en pourra être connue.

La promulgation faite par le PREMIER CONSUL sera réputée connue dans le département où siégera le Gouvernement, un jour après celui de la promulgation; et dans chacun des autres départemens, après l'expiration du même délai,

mètres [environ vingt lieues anciennes] entre la ville où la

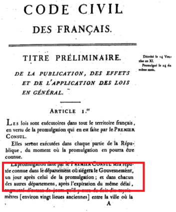
A

The standard CTD structure (Chisholm)

(1) *a* is obligatory
(2) If not-*a*, then *b* is obligatory
(3) If *a*, then not-*b* is obligatory

Benchmark problems

Group 1: CTD



CODE CIVIL
DES FRANÇAIS.

TITRE PRÉLIMINAIRE.

DE LA PUBLICATION, DES EFFETS ET DE L'APPLICATION DES LOIS EN GÉNÉRAL.

Décrété le 14 Ventôse an XI.
Promulgué le 24 du même mois.

ARTICLE 1.er

Les lois sont exécutoires dans tout le territoire français, en vertu de la promulgation qui en est faite par le PREMIER CONSUL.

Elles seront exécutées dans chaque partie de la République, du moment où la promulgation en pourra être connue.

La promulgation faite par le PREMIER CONSUL sera réputée connue dans le département où siégera le Gouvernement, un jour après celui de la promulgation; et dans chacun des autres départemens, après l'expiration du même délai, augmenté d'autant de jours qu'il y aura de fois dix myriamètres [environ vingt lieues anciennes] entre la ville où la

A

The standard CTD structure (Chisholm)

(1) *a* is obligatory

(2) If not-*a*, then *b* is obligatory

(3) If *a*, then not-*b* is obligatory

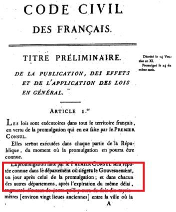
Primary obligation

CTD obligation

ATD obligation

Benchmark problems

Group 1: CTD



CODE CIVIL

DES FRANÇAIS.

TITRE PRÉLIMINAIRE.

DE LA PUBLICATION, DES EFFETS ET DE L'APPLICATION DES LOIS EN GÉNÉRAL.

Décrété le 14 Ventôse an XI. Promulgué le 24 du même mois.

ARTICLE 1.er

Les lois sont exécutoires dans tout le territoire français, en vertu de la promulgation qui en est faite par le PREMIER CONSUL.

Elles seront exécutées dans chaque partie de la République, du moment où la promulgation en pourra être connue.

La promulgation faite par le PREMIER CONSUL sera réputée connue dans le département où siégera le Gouvernement, un jour après celui de la promulgation; et dans chacun des autres départemens, après l'expiration du même délai, augmenté d'autant de jours qu'il y aura de fois dix myriamètres [environ vingt lieues anciennes] entre la ville où la

A

In the old days: SDL

KD modal logic

Obligatory : true in the ideal worlds

The standard CTD structure (Chisholm)

(1) *a* is obligatory
(2) If not-*a*, then *b* is obligatory
(3) If *a*, then not-*b* is obligatory
(4) Not-*a*

Problem

Formalisation

- Inconsistent

CTDs in the GDPR

Example 1

- Personal data shall be processed lawfully (Art. 5). For example, the data subject must have given consent to the processing of his or her personal data for one or more specific purposes (Art. 6/1.a).
- If the personal data have been processed unlawfully (none of the requirements for a lawful processing applies), the controller has the obligation to erase the personal data in question without delay (Art. 17.d, right to be forgotten).

Benchmark problems

Group 2: conflicts (cf. [Goble, 2013])

Normative conflict

The agent ought to do each of several things, but cannot do them all

Goble, L. (2013).
Prima facie norms, normative conflicts, and dilemmas.
In Gabbay, D., Horty, J., Parent, X., van der Meyden, R., and van der Torre, L., editors, *Handbook of Deontic Logic and Normative Systems*, pages 241–352. College Publications, London. UK.

Group 2: conflicts (cf. [Goble, 2013])

Normative conflict

The agent ought to do each of several things, but cannot do them all

Strict, or simple, conflicts: $\bigcirc a, \bigcirc \neg a$
Binary conflicts: $\bigcirc a, \bigcirc b \vdash \neg(a \wedge b)$
N-ary conflicts (general form): $\bigcirc a_1, \ldots \bigcirc a_n \vdash \neg(a \wedge \ldots \wedge a_n)$

Goble, L. (2013).
Prima facie norms, normative conflicts, and dilemmas.
In Gabbay, D., Horty, J., Parent, X., van der Meyden, R., and van der Torre, L., editors, *Handbook of Deontic Logic and Normative Systems*, pages 241–352. College Publications, London. UK.

Group 2: conflicts (cf. [Goble, 2013])

Normative conflict

The agent ought to do each of several things, but cannot do them all

Conflicts across regulations are common-place:

Must the user expressly provide consent to be tracked?

- GDPR: yes
- New EU e-privacy directive : not always.

The reasoner must be able to detect such conflicts.

Goble, L. (2013).
Prima facie norms, normative conflicts, and dilemmas.
In Gabbay, D., Horty, J., Parent, X., van der Meyden, R., and van der Torre, L., editors, *Handbook of Deontic Logic and Normative Systems*, pages 241–352. College Publications, London. UK.

Group 2: conflicts (cf. [Goble, 2013])

Normative conflict

The agent ought to do each of several things, but cannot do them all

State of the art with the modal logic approach: either the logic is too strong or too weak:

- Too strong: $\bigcirc a, \bigcirc\neg a \vdash \bigcirc b$ (deontic explosion)
- Too weak: $\bigcirc(a \vee b), \bigcirc\neg a \nvdash \bigcirc b$

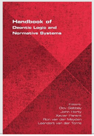
Goble, L. (2013).
Prima facie norms, normative conflicts, and dilemmas.
In Gabbay, D., Horty, J., Parent, X., van der Meyden, R., and van der Torre, L., editors, *Handbook of Deontic Logic and Normative Systems*, pages 241–352. College Publications, London. UK.

Our tool

I/O logic (in a nutshell)

One of the success stories of deontic logic

- Devised by Makinson & van der Torre
- Dedicated chapter in the **Handbook of Deontic Logic**



Conditionals (deontic reading): “If a, then b (obligation)”

- Semantics: “operational”
  - Procedures yielding outputs for inputs
- Proof-theory: generalizes existing ones
  - No axiom of identity (““If a, then a””)
  - Principle not desirable under a deontic reading

Our tool

A parenthesis on LegalRuleML

NOTEWORTHY

List of requirements (from the Oasis LegalRuleML TC)

Support for modelling:

1. Different types of rules
   - Prescriptive rules : obligations + permissions
   - Constitutive rules (counts-as conditionals)
     - Legal definitions ('sensitive personal data')
     - Legal interpretation
2. Defeasible reasoning (reasoning about exceptions)

I/O logic ticks all the boxes!

↪ Well-grounded in the legal domain

Our tool

A parenthesis on LegalRuleML

NOTEWORTHY

List of requirements (from the Oasis LegalRuleML TC)

Support for modelling:

1. Different types of rules
   - Prescriptive rules : obligations + permissions
   - Constitutive rules (counts-as conditionals)

     Legal definitions (‘sensitive personal data’)

     Legal interpretation
2. Defeasible reasoning (reasoning about exceptions)

I/O logic ticks all the boxes!

↪ Well-grounded in the legal domain

Our tool

Operational semantics

General form

Below: a conditional obligation in the I/O notation

body head

$\underbrace{(a, x)}_{\text{also called a `rule'}}$

a and x : two formulae of a base logic

A normative system N is a set of such pairs.

Main "semantical" construct:

$x \in O(N, a)$

Given input a (state of affairs), x (obligation) is in the output under norms N

Detachment (modus-ponens) : core mechanism of the semantics

Operational semantics

General form

Below: a conditional obligation in the I/O notation

body head

$$\underbrace{(a, x)}_{\text{also called a 'rule'}}$$

a and x : two formulae of a base logic

PL, FOL, modal logic, ...

A normative system N is a set of such pairs.

Main "semantical" construct:

$x \in O(N, a)$

Given input a (state of affairs), x (obligation) is in the output under norms N

Detachment (modus-ponens) : core mechanism of the semantics

Operational semantics

General form

Below: a conditional obligation in the I/O notation

body head

$$\underbrace{(a, x)}_{\text{also called a 'rule'}}$$

a and x : two formulae of a base logic

A normative system N is a set of such pairs.

Main "semantical" construct:

$x \in O(N, a)$??

Given input a (state of affairs), x (obligation) is in the output under norms N

Detachment (modus-ponens) : core mechanism of the semantics

Our tool

Two-sided architecture

Under the hood

On the outside

soundness and completeness results

$\Longleftarrow\!\Longrightarrow$

- Semantics

I/O operations

Proof-theory

Derivation tree

Nodes are pairs of formulas

Inference rules

$$N \longrightarrow (a, x)$$

But needed for, e.g., decidability
And hidden to the eyes of users

Clearing up a potential misunderstanding

"User-friendly" does not mean trivial.

When you see written on your screen

$$(a, x)$$

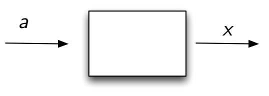
Clearing up a potential misunderstanding

"User-friendly" does not mean trivial.

When you see written on your screen

$$(a, x)$$

Under the hood: the reasoner has calculated that



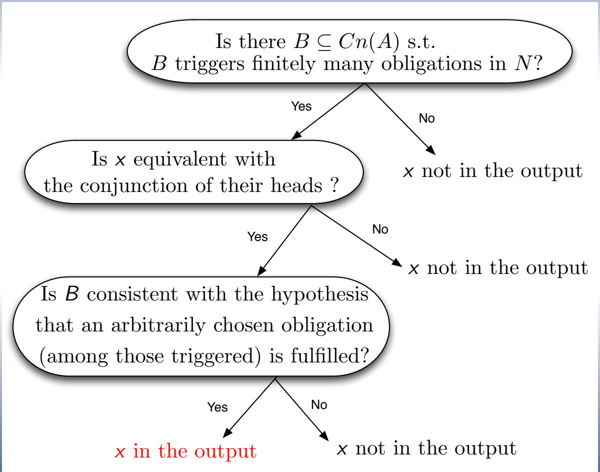
$$x \in O(N, a)$$

Semantics

How our solution works (roughly)

Our proposal: a new I/O operation

A bit involved: 3 steps (consistency check in the last)



Our take on the benchmarks (roughly)

Proof-theory

A modular treatment

Weaken the logic, but not too much

The troublemakers:

$$\text{WO}\ \frac{(a,x) \qquad x \vdash y}{(a,y)} \qquad\qquad \text{AND}\ \frac{(a,x) \qquad (a,y)}{(a, x \wedge y)}$$

It is okay to let WO go. It is not okay to let AND go.

Our take on the benchmarks (roughly)

Proof-theory

A modular treatment

Weaken the logic, but not too much

The troublemakers:

$$\text{WO}\ \frac{(a, x) \qquad x \vdash y}{(a, y)} \qquad\qquad \text{AND}\ \frac{(a, x) \qquad (a, y)}{(a, x \wedge y)}$$

It is okay to let WO go. It is not okay to let AND go.

Pragmatic oddity (Sergot/Prakken)

d: there is a dog
s: there is warning sign

$$\text{AND}\ \frac{\text{SI}\ \dfrac{(\top, \neg d)}{(d, \neg d)} \qquad (d, s)}{(d, \neg d \wedge s)}$$

No!

Our take on the benchmarks (roughly)

Proof-theory

A modular treatment

Weaken the logic, but not too much

The troublemakers:

$$\text{WO}\ \frac{(a, x) \qquad x \vdash y}{(a, y)} \qquad\qquad \text{AND}\ \frac{(a, x) \qquad (a, y)}{(a, x \wedge y)}$$

It is okay to let WO go. It is not okay to let AND go.

Conflicts

$$\frac{(a, x) \qquad (a, \neg x)}{(a, \bot)}$$

Yes!

$\bot$ is $x \wedge \neg x$
Inconsistencies across regulations

Our take on the benchmarks (roughly)

Proof-theory

A modular treatment

Weaken the logic, but not too much

The troublemakers:

$$\text{WO}\ \frac{(a, x) \qquad x \vdash y}{(a, y)} \qquad\qquad \text{AND}\ \frac{(a, x) \qquad (a, y)}{(a, x \wedge y)}$$

It is okay to let WO go. It is not okay to let AND go.

Reasoning about disjunction

$$\frac{(a, x \vee y) \qquad (a, \neg x)}{(a, \neg x \wedge y)}$$

Yes!

Our take on the benchmarks (roughly)

Proof-theory

A modular treatment

Weaken the logic, but not too much

The troublemakers:

$$\text{WO}\ \frac{(a,x) \qquad x \vdash y}{(a,y)} \qquad\qquad \text{AND}\ \frac{(a,x) \qquad (a,y)}{(a, x \wedge y)}$$

It is okay to let WO go. It is not okay to let AND go.

Proposal

$$\text{R-AND}\ \frac{(a,x) \qquad (a,y) \qquad a \wedge x \text{ consistent} \qquad a \wedge y \text{ consistent}}{(a, x \wedge y)}$$

Consistency checks

Summary

I have described a deontic reasoner currently under development.

- Well-defined:
  - Semantics and proof-theory
  - Completeness result linking the two
- Performs well on the two categories of benchmark problems of deontic logic

Spin-off: allows us to address a recurrent objection against rule-based systems

- Lack of proof-theory

Future work

A big step forward, but only only a first step.

Extensions

- Time, exceptions and other natural language constructs
- Permission and constitutive rules

Automation

- On-going work, with Benzmüller
  - Embedding into Higher-Order Logic (HOL)
  - Theorem-prover Isabelle/HOL for automation

Thank you!

Non-monotonic logic

Flavored by McCarthy and others.

Bottom line

$$\mathrm{SI}\ \frac{(\top, \neg d)}{(d, \neg d)}$$

Dashed line: the inference is blocked.

Sergot's view: no good for norm-compliance checking

Non-monotonic logic

Flavored by McCarthy and others.

Bottom line

$$\mathrm{SI}\ \frac{(\top, \neg d)}{(d, \neg d)}$$

Dashed line: the inference is blocked.

"The non-monotonic properties of a logic program using negation-by-failure make a consistent representation [of CTDs] possible. However, the program will have certain counter-intuitive properties. For instance, violated obligations simply vanish. Nothing more can be inferred about them, as the condition for something being obligatory no longer applies. One might argue that in actual life violated obligations do not vanish." (Herrestad)