An Introduction to the Philosophy of Logic

DANIEL COHNITZ
Utrecht University, the Netherlands

LUIS ESTRADA-GONZÁLEZ
National Autonomous University of Mexico

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107110939
DOI: 10.1017/9781316275573

First published 2019

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Cohnitz, Daniel, 1974– author. | Estrada-González, Luis, – author.
Title: An introduction to the philosophy of logic / Daniel Cohnitz, Utrecht University, the Netherlands, Luis Estrada-Gonzalez, National Autonomous University of Mexico.
Description: 1 [edition]. | New York : Cambridge University Press, 2019. | Includes bibliographical references and index.
Identifiers: LCCN 2018059147 | ISBN 9781107110939 (alk. paper)
Subjects: LCSH: Logic. Classification: LCC BC108 .C675 2019 | DDC 160–dc23
LC record available at https://lccn.loc.gov/2018059147

ISBN 978-1-107-11093-9 Hardback
ISBN 978-1-107-52772-0 Paperback

Contents

Figures

Tables

Preface and Acknowledgements

A lonely ship on the open ocean, as on the cover of this book, serves in several ways as metaphor in the philosophy of logic. Carnap, in his *The Logical Syntax of Language*, speaks of the "ship of logic" sailing on the "boundless ocean of unlimited possibilities" when introducing his pluralist conception of logic (Carnap, 1937), and Otto Neurath compared scientists in their attempts to revise their theories to "sailors who on the open sea must reconstruct their ship but are never able to start afresh from the bottom" (Neurath, 1973), a metaphor that is particularly apt when revisions of logic are concerned.

But the metaphor also holds in other ways. With the exception of Susan Haack's classic *Philosophy of Logics* (Haack, 1978), the philosophy of logic is largely uncharted open waters with many areas still labelled 'here be monsters'. We tried to provide a map for the fearless seafarer but we admit that we sometimes felt ourselves quite disoriented on this vast, deep ocean. We were glad that we could count on the help of various trawlers, shipmates, divers, pirates and other sailors whenever we felt lost. In alphabetical order, these were Francesco Berto, Rosalie Iemhoff, Annemarie Kalis, Johannes Korbmacher, Menno Lievers, Janneke van Lith, Niels van Miltenburg, Jesse Mulder, Carlo Nicolai, Juho Ritola, Marcus Rossberg, Erik Stei, Florian Gib Steinberger, Thomas Sturm, Albert Visser, the students of our course at the University of Tartu in 2013 and the research master students of Utrecht University in my course in 2016. When Luis wasn't available for completing the final stage of the book, we were glad that Shay Logan was ready to step in. He helped with Chapters 1 and 9 and did almost all the work for Chapter 2. We are very grateful for his help.

Introduction

The typical compulsory introduction to logic is odd: it's an odd course in a humanities curriculum, and it's particularly odd in a philosophy curriculum. What makes it somewhat odd in the humanities curriculum is the fact that it introduces a mathematical theory and methodology without much discussion of its scope and limits. There aren't many topics in the humanities that have been subject to rigorous mathematical analysis, and in the few cases in which scientists attempted to develop a mathematical model of their subject matter, the models are highly contested. However, logic courses seem unashamedly formal with little or no discussion of the approach's limits.

Within a philosophy curriculum, this situation is even stranger. Logic is supposed to be a universal methodology for doing philosophy, and logic is also a philosophical subdiscipline which has its origins in the work of Aristotle. Philosophy is a discipline that takes nothing for granted, and in which students are trained to develop a critical attitude towards any kind of claim. Nevertheless, in introductions to logic there is little room for argument, criticism or debate. Unlike pretty much any other course in the philosophy curriculum, logic is taught like a mathematics course.

It's not as if there weren't good reasons for doing it that way. Logic is one of the few subdisciplines of philosophy in which there has been considerable progress and convergence over the past 150 years. Much of this is due to the fact that – largely thanks to Gottlob Frege – modern logic is a rich mathematical theory. Introducing that theory and training students in it to a degree that enables them to apply the formal apparatus when analysing arguments for their validity, etc. is quite time consuming. Also, some of the most interesting questions about the scope and limits of logic can only be properly assessed when one knows at least a little bit about what logic actually is. Thus, it is well justified to postpone the critical discussion of logic for later and to limit introductions to logic to somewhat uncritical training courses in a formal methodology.

Nevertheless, this is unsatisfactory for the student of logic. She, for sure, would like to learn more about this discipline's philosophical foundations. But courses in philosophy of logic are unfortunately rare, witnessed or perhaps partly caused by the fact that there are almost no contemporary textbooks for such courses. This book is our attempt to fill that lacuna and provide a contemporary introduction to the central questions in the philosophy of logic.

What is the Philosophy of Logic?

This book is not an introduction to logic and it's not an introduction to philosophical logic either. As we will explain in the next chapter, philosophical logics are mathematical theories with a specific intended interpretation and application. This book presupposes acquaintance with an introduction to (philosophical) logic, as it is standardly taught in undergraduate programmes in philosophy all over the world. In particular, we assume a certain amount of familiarity with propositional logic and first-order (polyadic)[1] predicate logic. A good introduction to standard logic is Halbach (2010).

This book is also not intended as an introduction to the metatheory of standard logic or to non-classical deviations from standard logic. We will not presuppose acquaintance with metalogical results, or with deviant logics (and will try to introduce their gist when relevant), but this is not the place to discuss such matters in any (formal) detail. Good introductions to metalogic or deviant logic can be found instead in Sider (2010) and Priest (2008).

What the philosophy of logic is concerned with are the philosophical questions that relate to logic. For example, in the introductory course

[1] A predicate logic is "polyadic", if it contains in addition to one-place predicates also n-place relations. Thus, if you encountered formulas like '$\forall x \forall y \exists z(Rxy \rightarrow Rxz)$' in addition to formulas like '$\forall x(Fx \rightarrow Gx)$', you probably studied polyadic predicate logic. Predicate logic with only one-place predicates is called "monadic".

you learned to work with one specific logic, so-called "standard" or "classical" logic. There are alternatives to this logic: so-called "deviant" or "non-classical" logics. Why are there such alternatives? Is there reason to be discontent with classical logic? Is only one logic correct, or could several of these turn out to be correct? What does the correctness of a logic consist in? Getting the logical facts right? What are these facts? Are these mind-independent facts, are these facts about languages or conventions? How can we have knowledge of these facts? Do we know what's true in logic *a priori*? Could we be mistaken in our beliefs about logic? Could we detect such mistakes? Could we revise our beliefs about logic?

These and other questions will be discussed in this book (a detailed overview is in the next subsection). These questions concern the metaphysics, the epistemology and the methodology of logic. Some of these questions are similar to questions in related philosophical disciplines, such as the philosophy of language or the philosophy of mathematics. And if you are already familiar with these fields, you will recognize some of the arguments and moves from there. However, the philosophy of logic is a subdiscipline of philosophy by itself and not to be subsumed under either philosophy of language or philosophy of mathematics. That the questions that arise for logic are often unique is something that we hope you will learn from this book.

Overview of the Contents of this Book

Then let's dive into an overview of the topics we cover in this book. Of course, we couldn't cover every topic in the philosophy of logic, not even at the very introductory level this book is at. In our selection of material, we have tried to identify topics that we think are either the central questions of the philosophy of logic, or else closely related to them and in the centre of the contemporary discussion. For example, we haven't included a discussion of theories of truth or the reference of proper names in this book, because we didn't consider these topics to belong to the philosophy of logic (or at least not to its central questions). We also haven't included a detailed discussion of the philosophical controversy over the status of quantified modal logic, or the case for quantum logic, even though these topics belong to philosophy of logic. We felt that the relevant issues in these, the supposed metaphysical neutrality of logic, and the empirical revisability of logic, can be discussed in their own right, abstracting away from the specific logics.

We have divided the topics that we selected for inclusion into 10 somewhat equal-sized chapters.

1 The Nature and Tools of Logic

This chapter introduces the core terminology we will use in the book. What are we talking about when we talk about "logic"? Are we talking about certain mathematical theories or structures, or about a discipline? Can we make a distinction between *pure* and *applied* logic, in the way in which we can make that distinction for geometry? We also introduce a little bit of technical apparatus, borrowed from set theory, which we use for some of the more technical examples throughout the book. Most of these technicalities will be familiar to the reader, but we go over them to make sure that we are on the same page in terminology and notation. Finally, we present a general framework to talk about logic from a model-theoretic and a proof-theoretic point of view. Both perspectives will be relevant in later chapters.

2 The Standard Story and its Rivals

The second chapter introduces the main features of standard logic, reviews some arguments for and against them and presents some of the alternatives to standard logic. Some of the features of standard logic are so common that they might seem essential to logic. As we shall see in this chapter, but also in many of the later chapters, most of these features have been challenged.

3 Is Second-Order Logic Proper Logic?

In this chapter we continue to investigate deviations from standard logic. Here we will focus on the case of second-order logic. We sketch the standard and the Henkin semantics of second-order logic, and extend our proof system with rules for the second-order quantifiers. This chapter explains the differences between first-order logic with identity and second-order logic and the notion of categoricity. In many respects second-order logic seems just like a straightforward extension of standard logic, but it does come with some strange properties. Does that matter for whether second-order logic is logic? What makes a formal system a logic?

4 Logical Constants

What is the meaning of a logical connective? How is that meaning determined, and does it remain fixed across different logics? We present different ways of spelling out what the meaning of a logical constant is, Quine's classical argument for meaning-variance across logics and some attempts to block it.

What is usually regarded as the *problem of logical constants* is that of giving a principled demarcation between expressions that determine the logical form of an argument and those that do not. In the second half of Chapter 4, we discuss this problem, its motivation, potential solutions and some possible stances on its importance.

5 The Metaphysics of Logic

Logical realism is the view that logic is based on facts that are independent of us, our psychological make-up, inferential practices or conventions. Logical realism is perhaps a default view about logic; however, there are important alternatives. Psychologism about logic is the view that facts about our mind/brains ground logic. A further alternative to full-blooded realism and plain psychologism is to ground the logical facts in our rationality. This is a form of Kantianism about logic, recently defended by authors like Robert Hanna (2006a).

Another way of avoiding the idea that logic is made true by some fundamental features of reality, rationality or our minds is to hold that logical truths are "true by convention". This view was famously held by the logical positivists about both the laws of logic and those of mathematics. In Chapter 5 we will discuss the different options and introduce some of the main arguments of this debate about the metaphysical status of logic.

6 The Epistemology of Logic

If we follow the distinction between *logica docens*, *logica utens* and *logica ens* (as introduced in the first chapter), then we can distinguish three different questions about the revisability of logic. The first is the question of whether our beliefs about logic and our theories of logic can be revised, and what evidence there might be that could motivate us to such revision. The question of whether the logic we use is also revisable is already harder to answer. At least it seems as if training in (a particular) logic can influence our ways of reasoning. But can we rationally revise logic? Wouldn't that require to have a view from nowhere, which is a point of view that is impossible to take, since we always reason *in* a logic? Perhaps encountering problems within your own logic, for example by facing paradoxes that seem unacceptable by our own lights, might convince us that our logic needs revision. But how can we rationally evaluate the alternatives?

Closely related to the question of how we can rationally revise the logic we use is the question of how we can *justify* the logic we endorse. In this chapter we will explore what the problem of rule-circularity is and discuss possible answers to it: for example Nelson Goodman's suggestion that the laws of logic are justified because they are in reflective equilibrium with our judgements about instances of rule-application and the idea that we are default justified in applying the rules of logic and can hence escape the circle (or, rather, its viciousness).

7 Logical Pluralism

Plurality and logic may meet in different ways, from the now relatively uncontroversial 'There are many pure logics' to the highly controversial 'Some domains require different canonically applied logics' and 'There can be more than one correct logic canonically applied even to the same domain'. We probe these combinations and rank them according to their *prima facie* plausibility.

Perhaps the most interesting version of logical pluralism would be one in which logics are canonically applied to the same domain. In this chapter we explore the possible criteria that this kind of logical pluralism should meet to be (not trivially) true, or at least, to not be at a disadvantage vis-à-vis logical monism.

8 Logic, Reasoning and Rationality

If logic is a descriptive theory of (perhaps) the laws of truth or (perhaps) some very general features of metaphysical reality then we probably *should* reason along the laws of logic in the same sense as we should reason along the laws of physics when reasoning about physical subject matter.

However, there is a second way in which logic is often taken to be normative for thought. In this second way, logic is taken to be normative in a direct way, by understanding the laws of logic as telling us normatively how we ought to reason. As Gilbert Harman (1986) and other authors have pointed out, this doesn't seem right. If we understand logic in this way, then it would command us to conclude every arbitrary thing if we believed a single contradiction; and it would in any case clutter our belief box with all logical consequences of our beliefs, even though most of these consequences have no relevance to us whatsoever.

In the second half of Chapter 8 we will look at some of the puzzles that logic and our logical knowledge give rise to in epistemic logic and the semantics of propositional attitude reports. Perhaps logic is *a priori*, but certainly that doesn't mean that everyone *knows* all logical truths, right? Are there ways to distinguish different levels of logical knowledge? How can we characterize the increase in understanding as we proceed in proving a logical deduction?

9 Beyond Truth-Preservation

Some authors have argued that the most immediate justification of the idea that logical consequence preserves truth (from premises to conclusions) requires all the resources that yield a version of Curry's Paradox, so it could not be coherently required that valid arguments preserve truth. We survey here different arguments for and against the abandonment of truth-preservation based on such a paradox of validity. We also explore ways of logically relating premises and conclusions other than truth-preservation, that are mainly derived from the use of cognitive states like acceptance and rejection (or their linguistic expressions, assertion and denial) in the definition of validity.

10 The Place of Logic in Science

Perhaps the most evident role of logic in the empirical sciences is that of an inferential device used to extract information from certain bodies of beliefs or knowledge. However, there are other significant roles of logic in the sciences. Logic plays an important role in theoretical linguistics (for example, in the form of formal semantics), and as a theory of reasoning in cognitive science.

Mathematics is the paradigmatic case of logic working as an inferential device. But besides the problem of identifying a logic for mathematics given that seemingly one can do mathematics based on the logic one wishes, there is the problem of whether logic should be more properly thought of as a branch of mathematics, or as based on it, rather than the other way around.

In philosophy, formal logic is used as a tool. We use logic (and its symbolism) for disambiguation, the formal reconstruction of arguments, and for modelling. Especially in the last two functions, it seems that logic needs to satisfy special adequacy requirements: it must be philosophically neutral, and it must allow philosophers to express the kind of claims they intend to make.

Finally, we look at logic as a science itself. Is logic an exception in the family of sciences? If so, what makes it special? If not, how does it fit in?

How to Read the Book

As we said before, we presuppose familiarity with the contents of a standard introductory course in logic. It's possible to read the book to some extent without this familiarity, but we doubt that you will gain much insight from that. If you are not familiar with standard logic, we recommend reading Halbach (2010). This is a contemporary introduction to logic which uses similar terminology to ours.

The book is conceived as a course that can be read from beginning to end. However, the chapters are actually for the most part self-contained (and otherwise contain explicit references to earlier chapters). Chapter 1 introduces some of the core terminology we use, and Chapter 2 gives a quick recap about some of the main properties of standard logic. If you have read these two chapters, you should be able to jump to whatever topic interests you.

When using this book for self-study, we recommend to think about the questions at the end of the chapters and formulate answers to them. Even without external feedback this will probably give you an indication of how well you have understood the chapter.

How to Teach the Book

The course is conceived as a 10-week course with a lecture based on the textbook, and a seminar based on further readings, covering one chapter per week. If you have less time, you can concentrate on the textbook only and take the questions at the end of each chapter as starting points for discussion in the classroom. As we said above, Chapters 1 and 2 introduce and explain some of the terminology that is used and presupposed in other chapters. Other than that, you can freely pick and choose from the other chapters what you want to cover in your course.

1 The Nature and Tools of Logic

Before we can begin the discussion of the philosophical issues that logic gives rise to, we will first need to introduce a conceptual framework that will help us to disambiguate different senses of the term 'logic', and a bit of technical machinery that will allow us to talk about specific features and aspects of logic that play a role in the philosophical discussion of logic.

What is "Logic"?

The word 'logic' is used in many ways, it is a multiply ambiguous word. Note, for example, that one studies "logic" at college, but whereas one person may study Frege–Russell first-order logic, another person may study Aristotelian syllogistic or perhaps "informal logic and critical thinking". And that is not yet all, as it is not only that one may study different things under the label 'logic', but one also finds "logic gates" in circuit design or "fuzzy logic washing machines". Were homophony a sure sign of identity, we could conclude that logic is very powerful indeed, as it tortures students, sanctions philosophers' arguments and contributes to washing the dirty laundry of all of them.

However, the term 'logic' has usually been employed more or less indistinctly to designate a theory as well as what the theory is about and, moreover, also for naming several of its applications. For many disciplines, these aspects are linguistically distinguished, although the distinctions range from a few letters to different words. Consider the case of history where there is also often no clear distinction made. There is a widespread sense in which 'history' is used to designate a series of events in the past, and there is another equally widespread sense in which 'history' designates the science which studies those events, the former being thus the object of study of the latter. History is the study of history.

In the case of history, people use the same word, making only, but not always, a typographical differentiation. In the case of logic, people have usually made not even such differentiation, typographical or other. On the other hand, in the case of biology or physics there is a clear difference. Biology is the science of living beings (and not the study of biology), and physics is the science of matter and its transformations (and not the study of physics).

One might conjecture that the fact that this difference between the disciplines (biology and physics on the one side, logic and perhaps mathematics on the other) is due to the different degree of *reflexivity* of disciplines: the more reflexive a discipline is, the more blurred the distinction between the theory and its object. Consider the distinction between the *body of knowledge* of a discipline *D* and the *images* of *D*. The body of knowledge includes statements that are answers to questions concerning the object of study of *D*, as well as theories, facts, methods and open problems. Images, on the other hand, are cognitive and normative statements about the discipline that serve as guiding principles for both posing and answering questions that have arisen within the body of knowledge and which are typically not part of the body of knowledge. Such statements may determine which problems should be considered more pertinent and urgent, what counts as a pertinent experiment (or proof or argument), etc. The image also contains normative views on which procedures, individuals or institutions have the authority to settle disagreements inside the discipline and what the appropriate curriculum to educate the next generation of the members of the discipline should be.

On the basis of this distinction, Corry (2004) argues that mathematics is *reflexive*, in the sense that it can formulate and prove statements about itself from within the mathematical body of knowledge; i.e. certain parts of the body of knowledge of mathematics can be images of mathematics, and vice versa. So, could it be the high degree of reflexivity in mathematics and possibly logic that accounts for the difficulty in keeping theory and the object of study clearly separated?

Reflexivity in this sense is certainly an interesting notion, and studying its possible application in the case of logic would be worthwhile. However, it seems of little help for explaining the confusion between theory and object in disciplines like logic. One can well agree with Hilbert, Shapiro and many others that metamathematics is just more mathematics, and with

Corry on the reflexive character of mathematics, but there is no confusion between mathematics, its objects of study and metamathematics: Even if the objects of some mathematical theories could be other mathematical theories, this does not amount to an identification of theory and object; the levels still remain well distinguished.

Logic: Theory and Object – Pure and Applied

At any rate, once one is aware of the possibility of such confusion, one can better spend one's energy on developing mechanisms for avoiding those mistakes and confusions rather than on seeking the possible sources of the problems. Again, a comparison with certain parts of mathematics could be useful. Even for theorists who favour one or just a small selection of logics as the right ones, it is an undeniable fact that there are many logics. As we shall have occasion to see in the chapters that follow, there is standard logic, there are intuitionistic and constructive logics, relevance logics, paraconsistent logics, free logics and many more.

Now consider the analogy of geometry. In geometry, a domain of geometrical objects, such as points and curves, can be characterized by means of various principles, and we can study these objects for their mathematical interest. In this way, one can find Euclidean geometries that satisfy certain traditional assumptions about geometrical objects – for example, that the internal angles of a triangle total 180°. However, one can also find non-Euclidean geometries that do not satisfy these traditional assumptions – for example, the internal angles of a triangle might total more or less than 180°. Until the nineteenth century, 'geometry' just meant Euclidean geometry, but by the middle of that century the geometricality of non-Euclidean geometries was controversially discussed. Although at the time these alternative geometries weren't considered descriptions of the structure of physical space, they nevertheless appeared to deal with at least some sort of lines, points and other objects that seemed analogous to those of traditional Euclidean geometry.

Riemann realized that one might ask whether any of those geometrical theories, if interpreted accordingly, could apply to the physical world, whether there is anything like the actual geometry of the universe. Mere deduction of consequences from the principles does not suffice for establishing this, as suitable empirical interpretations of such principles are needed in order to apply the geometries to a not so pristine world. The correctness of a geometry seems to a great extent to be an empirical question. Almost fifty years after Riemann's work, the General Theory of Relativity postulated a connection between mass and the curvature of space(–time) which implied that space may have a non-zero curvature, and thus be non-Euclidean. Predictions of General Relativity were borne out by experimentation and the theory is now generally accepted as correct.

These developments in geometry led to a crucial distinction, namely that between *pure* and *applied* geometry in particular, and between pure and applied mathematics in general. Euclidean and non-Euclidean geometries, although mutually incompatible, can be studied on equal footing as pure mathematical theories. Can the situation in logic be tackled in a similar fashion, making sense of both the plurality of logics and logic's dual nature as theory and object?

Even though a distinction between pure and non-pure or applied logic has been advocated by several authors (see da Costa, 1979; da Costa, 1997; Bueno, 2001; Priest, 2003), the way the distinction is drawn varies. For simplicity, and because it has more unifying power, we will use a distinction akin to Graham Priest's version. Priest also stresses the distinction between logic as a theory and logic as an object of study (Priest, 2003, 207) or, perhaps better, he distinguishes the logical phenomena studied by logic as theory from the theory itself. Moreover, Priest has also introduced the fruitful notion of a *canonical application* of a theory. This conceptual stock will provide us with the necessary distinctions that we will make use of throughout the book.

The distinction pure/non-pure applies to each of the members of the distinction theory/object. Thus, we not only have pure and non-pure (applied) theories, but also pure and non-pure objects or phenomena. Let us start with the latter distinction. A *pure logic* would be a kind of mathematical structure, a logical structure. (What specific kind of mathematical structures the logical structures are will be discussed later.) A *non-pure logic* is a phenomenon with a logical structure. This includes daily-life arguments, the flow of electricity in a circuit, closures on topological spaces, the functioning of a computer program, etc.

A *canonical non-pure logic* is a phenomenon traditionally regarded as belonging to the proper domain of logic, such as formal or informal arguments made by people. Although the flow of electricity has a logical

Table 1.1 *Logic as a non-pure theory*

	Pure	**Non-pure**
Object	mathematical structures	arguments, flow of electricity
Theory	theories about mathematical structures	philosophical logics (like Frege's *Begriffsschrift*)

structure, it is not a canonical object of study of logic and that is why it counts as a non-canonical logical phenomenon.

We started the chapter by mentioning that the term 'logic' has often been used without distinction to designate a theory as well as for designating what the theory is about. For example, 'logic' has been used to name a certain science, roughly the science of right reasoning, as well as for the presumed subject matter of such science, roughly the structure of norms that govern right reasoning.[1] Some terminology should mirror this distinction. 'Logic', capitalized, could be used for theories that study pure logics, 'applied logic' for the study of empirical phenomena with a logical structure, and 'canonically applied logic' for the study of canonical empirical phenomena with a logical structure. It has been common practice to reserve the term without adjectives ('geometry', 'arithmetic') for the pure part. Noting a distinction between object and theory prevents most of the discussed problems in the use of the term 'logic'. If there is some reticence in using the word 'logic' for any of the parts, the study of pure logics might be called 'Universal Logic', as Béziau (1995) has suggested. Most of the times in this book, 'logic' will designate canonically applied logic, unless otherwise specified, since most of the more interesting philosophical discussions surrounding logic belong to its canonical side, whether this concerns logic as a theory or as an object. Nonetheless, sometimes context will determine what sense of 'logic' is being used, without requiring recurrent specification.

These distinctions will allow us to clarify certain questions and to pose them in better terms. A recent example of the utility of the distinctions is Priest's discussion of the revisability of logic (as we shall see in Chapter 6). Assuming that we are discussing the canonical part of logic, what is

[1] Whether logic is indeed best characterized as concerning norms of right reasoning will be discussed later at several places in this book. Here this just serves as an example.

considered to be in need of revision? Is it a theory, or is it an application of the theory or rather the object studied by the theory? Another example of the clarifications allowed by these distinctions concerns the discussion of logical pluralism (to which we come back in Chapter 7). What kind of plurality is championed or rejected? A plurality of pure logical theories with their corresponding plurality of pure logical structures? It seems that most philosophers of logic have a situation in mind in which one has a single canonical logical phenomenon, say, validity in everyday argumentation, and many canonically applied logical theories to explain the phenomenon. Is it possible that all these theories are right with respect to that single phenomenon? The distinctions above can be very helpful when approaching such questions.

Yet, for the purposes of this chapter, our initial question concerning the nature and subject matter of logic can also be disambiguated with the help of the distinctions just introduced. There need not be any tension between the idea that logic is a mathematical theory studying certain mathematical structures, and saying that logic is a theory about the evaluation of arguments in ordinary language and the sciences. Logic, as a pure mathematical theory, does the former, but also logic, as a canonically applied theory, does the latter. Now, logic has traditionally been expected to serve as a theory useful for the analysis of the inferential relationships between premises and conclusions expressed in arguments we actually employ (be it in the sciences or in daily life). Then, in many if not most debates, logic is considered a canonically applied theory.

Traditionally, logic is an account of logical validity that has to exhibit and explain certain features:

Truth-preservation If an argument is logically valid then the truth of the premises guarantees the truth of the conclusion.

Necessity That the premises imply the conclusion holds of necessity.

Formality Logically valid arguments are so in virtue of their logical form.

Aprioricity Logic is knowable *a priori*.

Universality A logically valid argument is valid across all domains of enquiry.

Normativity Rejecting a logically valid argument is somehow irrational.

Although all of them seem *prima facie* sound, how to spell these features and their connections out is an open philosophical problem. A more or less well-known case is that of *formality*. MacFarlane (2000) has argued that the formality of logic can mean at least three different things: (i) logic provides constitutive norms for thought as such; (ii) logic is indifferent as to the particular identities of objects; (iii) logic abstracts entirely from the semantic content of thought; to which Beall and Restall (2006) add (iv) logic is schematic. *Aprioricity* is another good example of the unsettledness of these features. That logic is *a priori* can mean, for example, that (i) valid arguments are so for non-empirical reasons; (ii) we can come to know the logical validity of an argument via non-empirical reasons; (iii) logic serves to organize the rest of our conceptual inquiries; (iv) logic or a logically valid argument cannot be disproved by empirical reasons. Similar remarks can be made for all the other features. Nonetheless, most logicians and philosophers of logic have endorsed that at least a significant part of these features should accompany logic, although nowadays it is difficult to find a single author that endorses all of them.

Different versions of each of these features will be discussed at length in the forthcoming chapters: *formality* in Chapters 3 and 4; *aprioricity* in Chapters 5 and 6; *universality* in Chapter 7; *normativity* in Chapter 8, and *truth-preservation* and *necessity* in Chapter 9.

In the course of the history of logic, in all its forms, several formal tools to study and do logic have been developed. But before turning to review some of the formal tools to study and do logic, we would like to examine an issue that has been hardly ever discussed, namely how far the analogy between geometry and logic should be taken.

Rescher's Rejection of the Geometric Analogy

One of the most detailed assessments of the analogy is Nicholas Rescher's (Rescher, 1969), which is at the same time one of the most severe critiques of it. According to Rescher, epistemological and developmental similarities between logic and geometry are not sufficient for claiming an analogy in more important matters, namely their respective treatment of rivalry, the question about the very nature of the theories involved or the characterization of their subject matter. We are not going to discuss the whole of connections between logic and geometry, but will just focus on the issue whether and how a distinction between the pure and the applied can be drawn in logic. We will argue that Rescher's denial of the analogy and his rejection of the idea of a pure logic are grounded in certain historical errors regarding the development of geometry and on a dogmatic assumption about the nature of logic as an essentially canonically applied theory.

Rescher accepts a general analogy between geometry and logic, namely that in the same way as there are many geometries there are also many logics. But this would be all there is to the analogy. Rescher's arguments proceeds from the following premises:

(R1) In geometry it is taken for granted that there is a distinction between "pure" and "applied" (or "physical") geometry.

(R2) Formally, all geometries are "right", or none can be formally "wrong" (because they are consistent etc.), but there could be only one right physical geometry.

It should be clear that Rescher is taking 'logic' to mean 'logic as a theory', as he is discussing correctness and other properties of theories. Rescher's first argument against the geometric analogy aims at establishing that there is nothing like "pure logic" in the way in which (R1) says that there is pure geometry (Rescher, 1969, 217). According to Rescher this is so because every system deserving the name 'logic' must satisfy, among other things, the requirement of having an interpretation which is not only a mathematical model for an abstract calculus nor just anything satisfying some formal axioms, but an interpretation involving concepts such as that of *sentences* (or another linguistic entity), *inferences*, *arguments*, as well as the concepts of *meaning* or *truth* of sentences. Mere mathematical theories, such as pure logics would be, do not involve any concepts of that kind. Rescher thus discards the distinction between pure and applied logic on the basis of a traditional view of logic. Because according to such a view logic is *essentially applied*, logic is identified with its canonical application.

Rescher uses (R2) in his second argument against the geometric analogy. He argues that even if we suppose that there are pure logics that are all formally correct, it would still be false that there is just one right applied logic. Rescher thinks that there are many correct logics. Even the mere possibility of the correctness of more than one logic is enough for him to regard logic as being essentially different from geometry.

Rescher's ultimate argument for rejecting the geometric analogy is that in order to develop a logic, one necessarily employs a "presystematic logic machinery". Rescher draws a distinction between "systematized logic", which would be a special branch of knowledge, a theory in the usual sense, and "presystematic logic", which would be a general instrument for the realization of knowledge throughout all its branches. The presystematic logic machinery would be the "preexisting idea of what logic is" (Rescher, 1969, 231) which would determine the following "regulative principles" or, better, criteria of logicality: precision, exactness, economy, simplicity, coherence and consistency, but "[a]bove all, one must [. . .] stress the regulative ideal of by-and-large conformity to the key features of the presystematic practice, of 'saving the phenomena' that are involved in the presystematic practice" (Rescher, 1969, 228). Such regulative principles would serve as criteria of logicality because they "will play a key role throughout the range of diverse 'logics' and their employment will effectively condition our view of such systematizations" (Rescher, 1969, 224). Thus, systems of logic would be systematizations of the presystematic practices of reasoning, normatively regarded. In other words, Rescher claims that in order to develop a system of logic one needs the preexisting idea of what logic is, while for developing a geometry no "presystematic geometry" is needed (and according to him such presystematic geometry does not even exist). Then he says that this difference implies "that the nature of the choice between alternative systems in the two cases will in fact have to be quite different" (Rescher, 1969, 219).

Let us discuss the last two arguments first. Rescher's second argument is inconclusive since its main premise, that there can be many (canonically) applied logics but just one applied geometry, can be seriously doubted. (Canonically) applied logical monism – roughly, the thesis that there is only one correct logical theory about right reasoning – and applied geometric pluralism are far more defensible than Rescher believes. We will discuss monism more thoroughly in Chapter 7, but an argument for logical monism put forward by Priest is that "[e]ven if modes of legitimate inference do vary from domain to domain, there must be a common core determined by the intersection of all these" (Priest, 2003, 464). That intersection would be the correct logic because its laws would be valid in all domains, and independently of every domain. It is possible that this intersection is empty, "but I never heard a plausible argument to this effect" (Priest, 2003, 464f). Concerning geometric pluralism, it is often held that different geometries are appropriate for different contexts. For example, when we build a tower it is appropriate to use Euclidean geometry. But when we do surveying, it is appropriate to use spherical geometry. And when we do cosmology, it is appropriate to use Riemannian geometry. It is possible that the usage of different geometries in different regions of the universe is not only a matter of simplicity, but also a requirement of the structure of the universe; maybe there is no global geometry. Rescher's second argument does not make its point; there is more needed if one wanted to draw an essential distinction between logic and geometry based on the status of monism and pluralism in each case.

Priest tries to reject Rescher's third argument as follows:

> Rescher's observation [that the articulation of a logical system requires a presystematic logic whereas the articulation of a geometric system does not require a presystematic geometry] seems correct. But [...] it is difficult to see it as having significant import for the question. The formulation of a grammar for a language requires the employment of a metalanguage, and so a metagrammar. But this hardly entails that there cannot be rival grammars for a language, or that the question of which is correct is not a posteriori. The same could be true of logic. (Priest, 2003, 443)

However, it seems that Priest misses Rescher's point. Rescher is contending neither logical pluralism nor the existence of rivalry in his third argument. Remember the second argument: he thinks that in the case of geometry only one will be the correct one among rival geometries; in the case of logic there are rivals, genuine alternatives, but many of them may be right. Rescher tries to show that there is a fundamental disanalogy between logic and geometry, that they are radically different disciplines, not that there are no rival logics nor that the question of which logic is correct should be answered *a priori*.

Alberto Coffa provides an nice summary of the state of affairs with respect to geometry at the end of nineteenth century and the beginning of twentieth century:

> During the second half of the nineteenth century, through a process still awaiting explanation, the community of geometers reached the conclusion that all geometries were here to stay [...]. [A] community of scientists had agreed to accept in a not-merely-provisory way all the members of a set of mutually inconsistent theories about a certain domain [...]. It was now up to philosophers [...] to make epistemological sense of the mathematicians' attitude toward geometry [...]. The challenge was a difficult test for philosophers, a test which (sad to say) they all failed [...]. (Coffa, 1986, 8)

A few pages later he continues:

> For decades professional philosophers had remained largely unmoved by the new developments, watching them from afar or not at all [. . .]. As the trend toward formalism became stronger and more definite, however, some philosophers concluded that the noble science of geometry was taking too harsh a beating from its practitioners. Perhaps it was time to take a stand on their behalf. In 1899, philosophy and geometry finally stood in eyeball-to-eyeball confrontation. The issue was to determine what, exactly, was going on in the new geometry. What was going on, one might reckon, was that geometry was becoming less the science of space or space-time, and more the formal study of certain structures. Issues concerning the proper application of geometry to physics were being separated from the status of pure geometry, the branch of mathematics. (Coffa, 1986, 17)

These passages support two observations. First, the idea of a presystematic geometry is not as odd as Rescher makes it look. The distinction between pure geometry and physical geometry is very recent, and it also had detractors coming from the “pure side”. Until the mid-nineteenth century, mathematics was mostly governed by empirical or more or less intuitive considerations. It was viewed as a collection of exact observations about the physical universe. Most mathematical problems arose from physics; in fact, there was no clear separation between mathematics and physics. Proof was a helpful method for organizing facts and reducing the chance of errors, but each physical fact remained true by itself regardless of any proof. This pre-nineteenth-century viewpoint persists in many textbooks until now, because textbooks change slowly, and because a more sophisticated viewpoint may be beyond the content of a textbook. Prior to the nineteenth century Euclidean geometry was seen as the best known description of physical space. Some non-Euclidean axioms for geometry were also studied, but not taken seriously; they were considered works of fiction. Indeed, most early investigations of non-Euclidean axioms were carried out with the intention of proving those axioms wrong: mathematicians hoped to prove that Euclid’s fifth postulate was a consequence of other axioms, by showing that a denial of the parallel postulate would lead to a contradiction. All such attempts were unsuccessful – the denial of the parallel postulate merely led to odd conclusions, not to outright contradictions – though sometimes errors temporarily led mathematicians to believe that they had succeeded in producing a contradiction. The motivation behind

Beltrami's and Klein's attempts to find models for non-Euclidian geometries was to understand them on the basis of the intuitive, Euclidean geometry, though their research eventually supported the mathematical autonomy of geometry. Around the turn of the century, Poincaré and Hilbert each provided an explanation of geometry that took the discipline to be an implicit definition of its concepts: its terms could be applied to any system of objects that satisfied the required axioms. Each of these two mathematicians found vigorous opposition from a different logicist – Russell against Poincaré and Frege against Hilbert – who maintained the vanishing view that geometry is essentially concerned with physical space or spatial intuitions. (Even contemporary philosophers still speak as if there were such presystematic geometry: "A finite projective plane is not going to be used to model physical space, but it may be used to model something analogous to physical space" (Beall and Restall, 2000, 489).) What happened was that mathematicians and philosophers revised their presystematic idea of "geometry" (and of "algebra" etc.) removing any empirical content from it and replacing most constant elements by variables.

The second point to take from Coffa's remarks is that it seems that logic has been suffering a transformation similar to that of geometry as described in the preceding paragraph. Nevertheless, most logicians, philosophers and even some mathematicians are still resistant to the idea of regarding logic as the formal study of certain structures rather than the science of the evaluation of arguments. This leads us to the analysis of Rescher's first argument and the consideration of the pure/applied distinction in the case of logic.

Once one distinguishes between logic as theory and object (both in their canonical forms), the pure/non-pure distinction (again in the canonical form) comes almost naturally. Once it is in place one can do justice to both intuitions, namely that there are pure logics, like there are pure geometries, as well as Rescher's traditional ideas about the subject matter of logic. The questions about rivalry and correctness only makes sense if one holds that logics have some special job to do, in the same way pure geometries are not rivals and none can claim to be "the right one" until they are applied to, say, the study of physical space. But those special jobs, on which there can be rivalry and correctness, are what we have called the canonical application of a logical theory. Rescher's claim would then be that canonically applied logics are not pure, and one cannot but agree. The further step is, as in the case of geometry, to recognize that the structure one retains after removing the special job from canonically applied logic is still a logic in a sense.

Background on Tools

In this book we try to keep technicalities to a minimum. We only refer to technical results when they are (or seem to be) of significance for the philosophical discussion. However, in order to present these technical results and to motivate them, we will need to have a certain technical toolkit in place.

In this section we introduce some set-theoretic concepts and some notation that will be useful for the rest of the book, but especially for discussions of those conceptions of logic in which the notion of model is central, and introduce thus the notion of model or interpretation. We also motivate the idea of proof theory in general and of sequent calculi in particular. Finally, we introduce the main ideas underlying the proof-theoretic conceptions of logic.

The following is not an introduction to mathematical logic. We presuppose familiarity with mathematical logic. The following is to some extent a generalization of what one typically learns in an introduction to logic, and an introduction to our preferred style of notation (to make sure that the reader is "on the same page" as us).[2]

Logic Model-theoretically

First, we will need some set-theoretical concepts and notation. A *set* is a collection of things, called its *elements*, such that no other relations beyond their belonging to it, and the identity and difference between them, matters for determining the set. In particular, neither the ordering nor the number of times each member appears in it matter. If the number of times matters, we have a *multiset*; if the ordering also matters, we have a *list*. (We will assume that each multiset has only finitely many members, each of which occurs only finitely many times.) Thus, even though $x \neq y$, $\langle x, y, x\rangle$ is a different list from $\langle x, x, y\rangle$, but $[x, y, x]$ is the same multiset as $[x, x, y]$;

[2] Again, for an introduction to modern logic, we recommend Halbach (2010) and Sider (2010).

$[x, x, y]$ is a different multiset from $[x, y]$, but $\{x, x, y\}$ is the same set as $\{x, y\}$. (Note the different types of brackets to collect members of lists, multisets and sets.) We will use 'object' to refer indistinctly to either a set, a multiset or a list. We use '$x \in X$' to say that x is an element of, or that it belongs to, the object X. We use '$X \cup Y$' for the union of X and Y – that is, for the set containing all and only those objects that are either in X or in Y. We use '$X \cap Y$' for the intersection of X and Y – that is, for the set containing all and only those objects that are both in X and in Y. Finally, we use $\emptyset$ for the empty set – that is, for the (unique!) set with no members.

A *map* is a relation between two objects X and Y – in that order, called the *domain* and *codomain* of the map – and denoted by '$f : X \longrightarrow Y$', such that to any element x in the domain corresponds one and only one in the codomain, often denoted '$f(x)$'.

Lists have to preserve the ordering between the elements, that is, if $\leq$ is the ordering between the elements of a list X, and a and b belong to it and are such that $a \leq b$, then $f(a) \leq f(b)$. Although domain and codomain can be different kinds of objects, say, the domain can be a multiset and the codomain a list, we will restrict ourselves to the cases where either both are of the same type or, if they are different, the codomain is a set. When we are just considering sets, a map will be called a 'function'.

Y^X is the object of all maps from X to Y, also called 'exponential'. The cardinality of Y^X, $|Y^X|$, can be computed by $|Y|^{|X|}$. In a typical first logic course, an interpretation of a formal language L is a map from L to $\mathcal{V}$, the set of (two) truth-values. Then, the size of a truth-table from your first logic course is an exponential of the form $\mathcal{V}^{\hat{L}}$, where $\hat{L}$ is a restriction of L, and it has as elements all the different propositional variables of the formula one is going to evaluate. So the cardinality of $\mathcal{V}^{\hat{L}}$ is $|\mathcal{V}|^{|\hat{L}|} = 2^n$, where n is the number of different propositional variables of the formula one is going to evaluate. 2^n thus gives the number of rows of a truth-table, that is, the number of interpretations of $\hat{L}$, that is, of elements of the exponential $\mathcal{V}^{\hat{L}}$.

$X \times Y$ is the object of all pairs $\langle x, y\rangle$ such that $x \in X$ and $y \in Y$. Again, in a typical first logic course, an n-ary connective k is interpreted as a map $f_k : \mathcal{V}^n \longrightarrow \mathcal{V}$, where '$\mathcal{V}^n$' is a shortand for $\mathcal{V} \times \ldots \times \mathcal{V}$ n times. Thus, unary connectives, like negation, are interpreted as maps $f : \mathcal{V}^1 \longrightarrow \mathcal{V}$, that is, $f : \mathcal{V} \longrightarrow \mathcal{V}$, whereas binary connectives are interpreted as functions of the form $f : \mathcal{V}^2 \longrightarrow \mathcal{V}$, that is, $f : \mathcal{V} \times \mathcal{V} \longrightarrow \mathcal{V}$. The two columns on the left in a typical truth-table give the pair members of $\mathcal{V} \times \mathcal{V}$ (see Table 1.2. for an example).

Table 1.2 *Example of a typical truth-table for classical conjunction*

ϕ	ψ	$(\phi \wedge \psi)$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$

Now we can give the basics of a model-theoretic presentation of a logic. The tabular presentation of logics hides, so to speak, many of the components of the models used to characterize such logics but that are also present in logics without truth tabular presentations, like standard first-order logic, intuitionistic logic or the post-Kripke modal logics. Think, for example, of a truth-table from (zeroth order) standard logic. It involves a number of more or less obvious components, such as (functional) evaluations that assign to each formula one and only one of the exactly two truth-values available. It is not very obvious whether these evaluations take evaluation indices into account and why, or that, actually, the truth-values are ordered in a certain way, and it is far less obvious how these components contribute to the determination of what counts as logically valid. We will try to make these aspects more precise. The following is thus a very general framework that can be used to characterize logics. It is a bit more complex than the standard model-theoretic characterization you find in logic textbooks, but it makes explicit many of the features that you can "tweak" in order to arrive at an alternative (pure) logic.[3]

Model-theoretically, a logic can be characterized by a structure of the form $S_{\mathcal{L}} = \langle L, W, \mathcal{R}, \mathcal{D}, \mathcal{V}, \mathcal{K}, \mathcal{C}, \delta, \mathfrak{v} \rangle$, where the tuple consisting in $S_{\mathcal{L}}$ minus L is called a *model* for L. The rough explanation of each component is as follows.

L stands for our *formal language*. The formal languages we are going to consider here are non-empty sets of formulas recursively defined as usual on (non-empty, possibly infinite) fixed formal vocabularies whose building blocks are

- (non-empty, possibly infinite) sets of *terms* (i.e. *constants* $a_1, a_2, \ldots$, *variables* $x_1, x_2, \ldots$ and *functional terms* $f_1, f_2, \ldots$);

[3] In the rest of the book, we will typically use only parts of this framework. However, we believe it is useful to get a somewhat comprehensive picture here.

- (non-empty) sets of sets of n-ary predicates for each natural number n R^1_1, $R^1_2, \ldots, R^n_1, R^n_2, \ldots$;
- (non-empty, possibly infinite) sets of sets n-ary connectives, which include quantifiers for any order m (we have thus *negations* $\neg_1, \neg_2, \ldots$, *conjunctions* $\wedge_1, \wedge_2, \ldots$, *disjunctions* $\vee_1, \vee_2, \ldots$, *conditionals* $\rightarrow_1, \rightarrow_2, \ldots$, *biconditionals* $\equiv_1, \equiv_2, \ldots$, *necessities* $\Box_1, \Box_2, \ldots$, *possibilities* $\Diamond_1, \Diamond_2, \ldots$, *universal quantifiers* $\forall^1_1, \forall^1_2, \ldots \ \forall^m_1, \forall^m_2$, *particular* (or *existential*) *quantifiers* $\exists^1_1$, $\exists^1_2, \ldots, \exists^m_1, \exists^m_2, \ldots$, etc.).

'p', 'q', 'r',... will be used as metavariables for atomic formulas (formulas with no connectives); 'τ_1', 'τ_2',... will be used as metavariables for generic terms; 'x', 'y', 'z',... will be used as metavariables for generic variables; 'a', 'b', 'c',... will be used as metavariables for generic constants; 'f', 'g', 'h',... will be used as metavariables for generic functional terms; 'F', 'G', 'H',..., 'P', 'Q', 'R',... will be used as metavariables for generic predicates (arity will be clear from the construction of the formula); 'ϕ', 'ψ',... will be used as metavariables for generic formulas. Also, most of the languages will contain only one of each kind of connective, but some of them may have more than one negation, or more than one conjunction, conditional, and so on.

W contains the *indices of evaluation*. These are where formulas are *evaluated* or assigned a semantic value. When doing modal logic, it is standard to call indices of evaluation *possible worlds*. More generally, indices can be thought of as (possibly empty) lists of conditions on which formulas are evaluated.

$\mathcal{R}$ contains a (possibly empty) collection of relations among the indices. Their role can be explained as follows: sometimes the semantic value of a formula at an index is computed using the values its compounds get at other indices. In order to keep track of these connections, one needs n-ary relations among indices; these are the elements of $\mathcal{R}$. If we continue with the example of modal logic, it is standard to evaluate *modal* formulas using a binary accessibility relation. Thus, if we were defining a modal logic, we would expect $\mathcal{R}$ to contain one binary relation for each modality being considered.

$\mathcal{D}$, the next element, is the *domain of quantification* function. For each type of quantifier Q in L, D assigns an appropriate class of objects D^Q_w to each $w \in W$. We will, throughout, adopt the convention of writing D or D_w to pick out the *first-order* domain of quantification. Continuing with the example of first-order modal logic, D is the function that assigns to each possible world the class of objects that exist in that world.

$\mathcal{V}$ contains the set of *truth-values*. It is standard to assume these are somehow ordered, and it is commonly assumed that $\mathcal{V}$ contains at least a value called *false*, denoted '$\perp$', and a value called *true*, denoted '$\top$', and that $\perp$ is not above $\top$.

$\mathcal{K}$ is a family of subsets of $\wp(\mathcal{V})$.[4] We take each element of $\mathcal{K}$ to be a *kind* of semantic value. As we shall see, $\mathcal{K}$ plays a role in defining different notions of logical consequence or validity. We will generally assume that $\mathcal{V}$ contains at least a set of *designated values*, written D^+ and a set of *antidesignated values*, written D^-. We will also assume the set of *designated values* satisfies the following conditions:

(a^+) $\top \in D^+$;
(b^+) for every $x, y \in \mathcal{V}$, if $x \in D^+$ and $y \notin D^+$ then $x \nleq y$.

And we will also assume that the set of antidesignated values satisfies the following conditions:

(a^-) There is an $x \in \mathcal{V}$ such that for every $y \in \mathcal{V}$, $x \leq y$. Let us call *false* such x and denote it '$\perp$'. $\perp \in D^-$;
(b^-) for every $x, y \in \mathcal{V}$, if $x \in D^-$ and $y \notin D^-$ then $y \nleq x$.

$\mathcal{C}$ contains the required *truth relations*. Each of these is a relation between constructions on the set of truth-values and the set of truth-values, and each plays a role in the evaluation of complex formulas. To say more, it helps to consider the example of the usual connectives, about which more will be said in the next chapter.

To begin, recall that negations, necessities and possibilities are *unary* connectives; and that conjunctions, disjunctions, conditionals and biconditionals are *binary* connectives. All of these are said to be of zeroth order because they don't require quantification over individuals or sets of (sets of) them. Thus, generally speaking, for each unary connective u in L, there will be an element ρ_u in $\mathcal{C}$ of the form $\mathcal{V} \times W \longrightarrow \mathcal{V}$, while for each binary connective b in L, there will be an element ρ_b in $\mathcal{C}$ of the form $\rho_b : (\mathcal{V} \times \mathcal{V}) \times W \longrightarrow \mathcal{V}$.

[4] $\wp(\mathcal{V})$ is the powerset of $\mathcal{V}$, the set of all subsets of $\mathcal{V}$ (which includes the empty set and $\mathcal{V}$ itself).

Quantifiers will also be thought of as unary connectives, this time of order greater than zero. So quantifiers come in a variety of forms: first-, second-, third-order and so on. First-order quantifiers quantify over individuals, second-order quantifiers quantify over sets of individuals, third-order quantifiers quantify over sets of sets of individuals, and so on. Thus, corresponding to each first-order quantifier f in L there will be a function ρ_f of $\mathcal{C}$ of the form $\mathcal{V}^{\mathcal{D}} \times W \longrightarrow \mathcal{V}$; corresponding to each second-order quantifier s a function ρ_s of the form $\mathcal{V}^{\mathcal{V}^{\mathcal{D}}} \times W \longrightarrow \mathcal{V}$, and so on.

The next element, δ, is the *denotation function*. It recursively specifies the denotation of each term at each $w \in W$. It does this by first assigning an appropriate function $\delta_w(f)$ to each function symbol f, and assigning an appropriate element $\delta(t)$ to each term t. By 'appropriate' in each case we mean 'appropriate to the type of the term/function'. We also take δ to have assigned a denotation $\delta_w(t)$ to each term t at each $w \in W$ by the following recursive procedure:

- If t is a constant or variable, then $\delta_w(t)$ is whatever δ assigns to t.
- If f is an n-ary function symbol and $t_1, \ldots, t_n$ are terms, then $\delta_w(f(t_1, \ldots, t_n)) = \delta_w(f)(\delta_w(t_1), \ldots, \delta_w(t_n))$.

Finally, $\mathfrak{v}$ is a valuation function. It assigns a value in $\mathcal{V}$ at each $w \in W$ to any propositional parameters in L. It also assigns to each n-ary predicate P and index w, a function $\mathfrak{v}_w(P) : \mathcal{D}^n \to \mathcal{V}$. This suffices, together with the previously specified information, to recursively extend $\mathfrak{v}$ to a function from $L \times W$ to $\mathcal{V}$.

The following will serve as both an example and an assumption: we will assume in the remainder of the book that any language containing one of the following connectives will interpret it in the given way:

$$\mathfrak{v}_i(\neg\phi) = \rho_\neg(\mathfrak{v}_i(\phi)) = \begin{cases} \top & \text{if } \mathfrak{v}_i(\phi) = \bot \\ \bot & \text{if } \mathfrak{v}_i(\phi) = \top \end{cases}$$
$$\mathfrak{v}_i(\phi \wedge \psi) = \rho_\wedge(\mathfrak{v}_i(\phi), \mathfrak{v}_i(\psi)) = \min(\mathfrak{v}_i(\phi), \mathfrak{v}_i(\psi))$$
$$\mathfrak{v}_i(\phi \vee \psi) = \rho_\vee(\mathfrak{v}_i(\phi), \mathfrak{v}_i(\psi)) = \max(\mathfrak{v}_i(\phi), \mathfrak{v}_i(\psi))$$
$$\mathfrak{v}_i(\phi \rightarrow \psi) = \rho_\rightarrow(\mathfrak{v}_i(\phi), \mathfrak{v}_i(\psi)) = \max(\mathfrak{v}_i(\neg\phi), \mathfrak{v}_i(\psi))$$
$$\mathfrak{v}_i(\Box\phi) = \rho_\Box(\phi) = \inf\{\mathfrak{v}_j(\phi) : Rij\}$$
$$\mathfrak{v}_i(\Diamond\phi) = \rho_\Diamond(\phi) = \sup\{\mathfrak{v}_j(\phi) : Rij\}$$

There are probably a few things here – namely, the min, max, inf and sup operators – that are unfamiliar to many readers. We'll briefly explain what these mean; first, though, recall that we've assumed that the members of $\mathcal{V}$ are somehow ordered. Let's write $\sqsubseteq$ for this ordering, whatever it is. The basic idea is that, using this ordering, we can talk about least elements (mins), greatest elements (maxes), greatest lower bounds (infs), and least upper bounds (sups). More explicitly,

- $\min(a,b) = \begin{cases} a & \text{if } a \sqsubseteq b \\ b & \text{otherwise} \end{cases}$
- $\max(a,b) = \begin{cases} b & \text{if } a \sqsubseteq b \\ a & \text{otherwise} \end{cases}$
- If X is a subset of $\mathcal{V}$, then we say that $\inf(X) = a$ iff (if and only if)
  - (i) $a \sqsubseteq x$ for all $x \in X$, and
  - (ii) If $b \sqsubseteq x$ for all $x \in X$, then $b \sqsubseteq a$.
- If X is a subset of $\mathcal{V}$, then we say that $\sup(X) = a$ iff
  - (i) $x \sqsubseteq a$ for all $x \in X$, and
  - (ii) If $x \sqsubseteq b$ for all $x \in X$, then $a \sqsubseteq b$.

Types of Validity

So, model-theoretically, we take logics to be a type of 9-tuple. But how does this help us give an account of validity? We'll explain by way of example. First recall that one of the elements of our model-theoretically understood logics was a set $\mathcal{K}$ of subsets of our set of semantic values $\mathcal{V}$. We said these corresponded to *kinds* of semantic values. We also assumed that $\mathcal{K}$ would be taken to contain a set of *designated values* D^+ and a set of *antidesignated values*, D^-. Roughly speaking, the set of designated values is the set of *ways to be true*, while the set of antidesignated values is the set of *ways to be false*.

With that in mind, we might say that an argument is D^+*-logically valid*, or less precisely that it is *truth-preserving*, and in any case denoted $\Gamma \vDash^{D^+}_{\mathcal{L}} \psi$, if and only if whenever $\mathfrak{v}(\phi) \in D^+$, for all $\phi \in \Gamma$, $\mathfrak{v}(\psi) \in D^+$ too. This gives us one plausible account of validity – an argument is valid when it transmits truth from premises to conclusion.

Several related logical notions are nearby. For example, we might say a formula ϕ is *satisfiable* if and only if there is a model in which $\mathfrak{v}(\phi) \in D^+$ – that is, if and only if it *can be true* in some way. Or we might say a *set* of formulas Γ is satisfiable if and only if there is a model in which $\mathfrak{v}(\phi) \in D^+$ for every $\phi \in \Gamma$ – that is, if and only if they can all simultaneously be true in some way. We hope it's clear that we could go on.

On the other hand we might say that an inference is D^--valid in $\mathcal{L}$, denoted $\Gamma \vDash_{\mathcal{L}}^{D^-} \psi$, if and only if whenever $\mathfrak{v}(\psi) \in D^-$, $\mathfrak{v}(\phi) \in D^-$ for some $\phi \in \Gamma$ as well. Roughly, the idea is that ψ is a D^--logical consequence of Γ if and only if every model that fails to satisfy ψ is also a model that fails to satisfy Γ. In many cases, D^+- and D^--logical consequence coincide for a given language L, i.e. they determine the same collection of valid arguments. However, in Chapter 9 we will discuss some cases in which that is not so. Just as in the case of D^+-validity, there are a variety of related logical notions available in the case of D^--validity – but the details are much the same as above, so we will leave thinking them through to the reader.

Model Theory: Summing Up

All of this was rather abstract, and the reader could be forgiven for having skimmed or even skipped much of it. In the next chapter, we will be examining several concrete examples. It is our opinion that neither the concepts presented here nor the concepts presented there can truly be understood independently of each other. Rather, what one ought to do is read one, then the other, then the first again, etc. until one has a wholistic grasp of both chapters.

In any event, before we turn to examples, there is one topic left to cover in *this* chapter: proof theory.

Logic Proof-theoretically

In this book, when a logic, or certain features of it, need to be presented proof-theoretically, we will work with sequent calculi to do it, since this kind of presentation allows for an easy focus on the structural properties of a logic, which will be important for many of the topics discussed in the book.

The proof-theoretic framework provides the resources to analyse logical consequence as a procedural concept. In the proof-theoretic tradition, logical consequence is understood as something more epistemologically robust than mere preservation of values over a class of models, and the main object of study of logic is the stepwise right reasoning from premises to conclusions. Such an approach involves taking seriously the idea of looking at structural properties of proofs as an integral part of the concept of logical consequence. This is not necessarily confusing the question of what logical consequence is with how we come to know that a conclusion is a logical consequence of a set of premises. Proof-theorists think that any rewarding conceptual analysis of logical consequence must adhere to the idea that at the heart of logic is the act of inferring, not the property of preservation of certain values.

It is important to emphasize that the proof-theoretic approach to consequence is taken to be an independent enterprise; in particular, their adherents resist the idea that it has to be sound with respect to a model-theoretic consequence relation.[5] Just like model-theoretic consequence relations are somehow "sound" by internal standards, not with respect to some other formal relation, the proof-theoretic approach ought to give an account of consequence whose success is not measured by comparison to some other formalism, but by general desiderata of the conceptual analysis of logical consequence.

Now to the details of sequent calculi. A *sequent* is something of the form $\Gamma \Rightarrow \Delta$, where Γ and Δ are lists of formulas in a formal language like that mentioned in the previous subsection. Sequents stand in for arguments; Γ comprises the *premises* of the sequent, and Δ its *conclusions* (also called 'antecedent' and 'succedent' of the sequent, respectively).

A *sequent calculus* gives us rules for deriving sequents, and each calculus determines a set of derivable sequents, which we will call the *consequence relation* determined by the calculus. Intuitively speaking, the rules encode ways of transforming inferences in an acceptable way. These rules come in two varieties: operational and structural. An *operational rule* is one involving certain specific vocabulary; we will introduce a number of these rules later.

Structural rules, on the other hand, do not involve any specific vocabulary. They apply to any formulas, independent of their shape. There are five main structural rules we will consider here. Identity allows us to derive $A \Rightarrow A$ from nowhere, for any A; it guarantees a reflexive consequence relation. (It is often considered an axiom scheme rather than a rule; we will count it as a zero-premise rule for uniformity.) *Weakening* allows us to add whatever premises or conclusions we like to any sequent we have derived, and ensures that our consequence relations will be monotonic. *Exchange* allows us to disregard the order in which formulas appear in a sequent.

5 In this book we will denote a model-theoretic consequence relation with a double turnstyle, $\models$, but denote a proof-theoretic consequence relation with a single turnstyle, $\vdash$.

Contraction allows us to move from a sequent that uses a formula twice as a premise or a conclusion to the corresponding sequent that uses it only once. In the presence of Exchange, together with certain special cases of Weakening, we can replace lists with multisets in our sequents. In the presence of Contraction (together with certain special cases of Weakening), we can replace multisets with ordinary sets in our sequents. *Cut* allows us to combine derivations in a certain way: if we can derive A as a premise in a certain context, and A as a conclusion in a certain context, then we can put those contexts together without A, and derive the result. This is a generalization of *Transitivity*: when Cut holds, the resulting logic is transitive. (Starting from the premise-sequents $B \Rightarrow A$ and $A \Rightarrow C$, an application of Cut yields the conclusion-sequent $B \Rightarrow C$.)

Reflexivity

$$\frac{}{\phi \Rightarrow \phi}$$

Weakening

$$\text{WL}\ \frac{\Gamma \Rightarrow \Delta}{\phi, \Gamma \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \phi}\ \text{WR}$$

Exchange

$$\text{EL}\ \frac{\Gamma, \phi, \psi, \Delta \Rightarrow \Pi}{\Gamma, \psi, \phi, \Delta \Rightarrow \Pi} \qquad \frac{\Gamma \Rightarrow \Delta, \phi, \psi, \Pi}{\Gamma \Rightarrow \Delta, \psi, \phi, \Pi}\ \text{ER}$$

Contraction

$$\text{CL}\ \frac{\phi, \phi, \Gamma \Rightarrow \Delta}{\phi, \Gamma \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \phi, \phi}{\Gamma \Rightarrow \Delta, \phi}\ \text{CR}$$

Cut

$$\frac{\Gamma \Rightarrow \Delta, \phi \qquad \phi, \Pi \Rightarrow \Sigma}{\Gamma, \Pi \Rightarrow \Delta, \Sigma}$$

How to *read the rules* in a sequent system? In principle, the rule says that a sequent of the form below the line may be inferred from (a) sequent(s) of the form above the line. Our target logics have more than just structural rules, of course. There are also *operational rules*, rules that govern particular bits of vocabulary. We will keep it relatively simple here and focus on the usual connectives: negation, conjunction, disjunction, implication and universal and particular quantifiers. Even with this restricted scope, though, there is a lot to work with.

Negation

$$\neg L\ \frac{\Gamma \Rightarrow \Delta, \phi}{\neg\phi, \Gamma \Rightarrow \Delta} \qquad \frac{\phi, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg\phi}\ \neg R$$

$$\neg L^*\ \frac{\Gamma \Rightarrow \phi, \Delta}{\Gamma, \neg\phi, \Rightarrow \Delta} \qquad \frac{\Gamma, \phi \Rightarrow \Delta}{\Gamma \Rightarrow \neg\phi, \Delta}\ \neg R^*$$

Conjunction

$$\frac{\phi, \Gamma \Rightarrow \Delta}{\phi \wedge \psi, \Gamma \Rightarrow \Delta} \qquad \frac{\psi, \Gamma \Rightarrow \Delta}{\phi \wedge \psi, \Gamma \Rightarrow \Delta}\ \wedge L \qquad \frac{\Gamma \Rightarrow \Delta, \phi \qquad \Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \phi \wedge \psi}\ \wedge R$$

$$\frac{\phi, \psi, \Gamma \Rightarrow \Delta}{\phi \wedge \psi, \Gamma \Rightarrow \Delta}\ \wedge L^* \qquad \frac{\Gamma \Rightarrow \Delta, \phi \qquad \Pi \Rightarrow \Sigma, \psi}{\Gamma, \Pi \Rightarrow \Delta, \Sigma, \phi \wedge \psi}\ \wedge R^*$$

Disjunction

$$\frac{\phi, \Gamma \Rightarrow \Delta \qquad \psi, \Gamma \Rightarrow \Delta}{\phi \vee \psi, \Gamma \Rightarrow \Delta}\ \vee L \qquad \frac{\Gamma \Rightarrow \Delta, \phi}{\Gamma \Rightarrow \Delta, \phi \vee \psi} \qquad \frac{\Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \phi \vee \psi}\ \vee R$$

$$\frac{\phi, \Gamma \Rightarrow \Delta \qquad \psi, \Pi \Rightarrow \Sigma}{\phi \vee \psi, \Gamma, \Pi \Rightarrow \Delta, \Sigma}\ \vee L^* \qquad \frac{\Gamma \Rightarrow \Delta, \phi, \psi}{\Gamma \Rightarrow \Delta, \phi \vee \psi}\ \vee R^*$$

Conditional

$$\frac{\Gamma \Rightarrow \Delta, \phi \qquad \psi, \Pi \Rightarrow \Sigma}{\phi \rightarrow \psi, \Gamma, \Pi \Rightarrow \Delta, \Sigma}\ \rightarrow L \qquad \frac{\phi, \Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \phi \rightarrow \psi}\ \rightarrow R$$

$$\frac{\Gamma \Rightarrow \Delta, \phi \qquad \psi, \Gamma \Rightarrow \Delta}{\phi \rightarrow \psi, \Gamma \Rightarrow \Delta}\ \rightarrow L^* \qquad \frac{\phi, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \phi \rightarrow \psi} \qquad \frac{\phi, \Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \phi \rightarrow \psi}\ \rightarrow R^*$$

Universal quantifier

$$\forall L\ \frac{\phi(\tau), \Gamma \Rightarrow \Delta}{\forall x\phi(x), \Gamma \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \phi(a)}{\Gamma \Rightarrow \Delta, \forall x\phi(x)}\ \forall R^\dagger$$

Existential quantifier

$$\exists L^\dagger\ \frac{\phi(a), \Gamma \Rightarrow \Delta}{\exists x\phi(x), \Gamma \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \phi(\tau)}{\Gamma \Rightarrow \Delta, \exists x\phi(x)}\ \exists R$$

Notice we've marked both $\forall R$ and $\exists L$ with a '†'. This is because there is a restriction on these rules: in both cases a must be what is called an *eigenvariable*. In both cases this means the same thing: a must not occur freely anywhere in the bottom sequent in either of these inferences.

The starred pairs define what are usually called 'intensional', 'multiplicative' or 'group-theoretical' connectives, in contrast to the non-starred pairs, which are called 'extensional', 'additive' or 'lattice-theoretical' connectives.

Finally, all four negation rules are "double line" rules. This is a convenient shorthand; what it means is just that the rule one gets by turning any of these four rules upside-down is also a rule in our system.

A *proof* in a sequent calculus *SC* is a finite labelled tree whose nodes are labelled by sequents, in such a way that leaves are labelled by axioms and each sequent at a node is obtained from sequents at immediate predecessor nodes according to one of the rules of *SC*. We will denote proofs by means of the metavariables $D, D', \ldots$. If D is a proof, a subtree D' of D which is itself a proof is called a *subproof* of D. A sequent S is provable in *SC* (or *SC*-provable, or a theorem of *SC*) if and only if it labels the root of some proof in *SC* (i.e., as it is sometimes said, if and only if it is the endsequent of such a proof). In the next chapter, we will look at a slightly modified form of this system. For now it's worthwhile to see the system 'in action'.

To begin, let's think about how to prove $p \Rightarrow q \to p$. Since our goal is to show this sequent can occur as the endsequent of a proof, we begin by writing it down, then applying rules to it to grow the proof *upwards* from this root, eventually reaching leaves of the form $r \Rightarrow r$. In stages, one might do this as follows:

Stage 1: $p \Rightarrow q \to p$

Stage 2: $$\dfrac{p, q \Rightarrow p}{p \Rightarrow q \to p}\ {\to}\,\mathrm{R}$$

Stage 3: $$\dfrac{\dfrac{q, p \Rightarrow p}{p, q \Rightarrow p}\ \mathrm{EL}}{p \Rightarrow q \to p}\ {\to}\,\mathrm{R}$$

Stage 4: $$\dfrac{\dfrac{\dfrac{p \Rightarrow p}{q, p \Rightarrow p}\ \mathrm{WL}}{p, q \Rightarrow p}\ \mathrm{EL}}{p \Rightarrow q \to p}\ {\to}\,\mathrm{R}$$

Similarly, if we want to prove that $p \to q \Rightarrow \neg p \vee q$, then we might do so as follows:

Stage 1: $p \to q \Rightarrow \neg p \vee q$

Stage 2: $$\dfrac{p \to q \Rightarrow \neg p, q}{p \to q \Rightarrow \neg p \vee q}\ {\vee}\,\mathrm{R}^*$$

Stage 3: $$\dfrac{\dfrac{p \to q, p \Rightarrow q}{p \to q \Rightarrow \neg p, q}\ {\neg}\,\mathrm{R}^*}{p \to q \Rightarrow \neg p \vee q}\ {\vee}\,\mathrm{R}^*$$

Stage 4: $$\dfrac{\dfrac{\dfrac{p \Rightarrow p \qquad q \Rightarrow q}{p \to q, p \Rightarrow q}\ {\to}\,\mathrm{L}}{p \to q \Rightarrow \neg p, q}\ {\neg}\,\mathrm{R}^*}{p \to q \Rightarrow \neg p \vee q}\ {\vee}\,\mathrm{R}^*$$

Questions

1. Sort the following into the categories of Table 1.1: Aristotelian syllogistics, electrical circuits, entailment relations between propositions, abstract objects instantiating a boolean algebra.

2. Are all the features of logic that we listed on page 15 in your view *essential* to logic? Are there features that in your view do not hold or are more controversial than others? If so, why?
3. Consider whether it makes sense to expect that non-canonically applied logics exhibit some of the features listed on page 15. In what sense could logic gates in circuit design be normative? What is the normativity of a pure logic (or of group theory for that matter)?
4. *Formality* is often identified with *topic-neutrality*, but isn't a pure logical theory concerned with a particular subject matter, namely logical structures, and thus not topic-neutral? What is *topic-neutrality* anyway? Canonically applied logic has a specific subject matter and that does not seem to affect its topic-neutrality. How can an object, a logical structure, be topic-neutral?
5. To make sure you understand the difference between min and inf and between max and sup, answer the following questions:
   - As given, min and max work on only two elements at a time. How would you generalize min and max to make sense of expressions like $\min(x_1, x_2, \ldots, x_n)$ when $n > 2$?
   - Explain why $\min(x_1, x_2, \ldots, x_n) = \inf(x_1, x_2, \ldots, x_n)$.
   - Suppose $\mathcal{V}$ is identical to the set of rational numbers. Let X be the set of rational numbers strictly greater than 0 and strictly less than 1. Explain why $\min(X)$ and $\max(X)$ are undefined, but $\inf(X) = 0$ and $\sup(X) = 1$.
   - Be careful! One should not conclude from the previous part of the problem that we can't make sense of $\min(X)$ and $\max(X)$ whenever X is infinite. Show by example that there is a perfectly good way to make sense of these concepts.

2 The Standard Story and Its Rivals

Written by Shay Logan

Introduction

If there *is* a "standard story" about the nature of logic, then it has the following three pieces:

Definition of Logical Consequence (DLC) Δ is a logical consequence of Γ if and only if it is impossible for every sentence in Γ to be true while every sentence in Δ is not true.

Consistency Assumption (CN) It is impossible for a sentence to simultaneously be both true and false.

Completeness Assumption (CM) It is impossible for a sentence to be neither true nor false.

DLC is, you might see, essentially the definition of D^+-logical consequence from the previous chapter. It may not seem like adding CN and CM to DLC does very much. But, as it turns out, together these three theses are sufficient to single out a unique correct logic in many formal languages. The logics that arise from these assumptions have come to be called *classical logics*. This is not because of any connection to anything else that goes by the name 'classical' but because they are, as the logics that one gets from the Standard Story, something akin to standard logics.

In this chapter we'll see how to get from the Standard Story to classical logic in one particularly simple language. We'll then turn to examining alternatives that arise by challenging various of the assumptions in the Standard Story. Along the way we will examine philosophical arguments for and against these assumptions.

From the Standard Story to classical logic

To give a flavour for how the Standard Story gives rise to a unique logic, we will restrict our attention to a simple, non-modal propositional language. Attention to detail will turn out to pay dividends here, so we'll take the time to set things up very explicitly.

Linguistic Matters

We will call the language we are discussing '$\mathcal{L}$'. The complete vocabulary of $\mathcal{L}$ will consist of the following:

- the *propositional variables* $p, q, r, p_1, q_1, \ldots,$
- the *unary connective* $\neg$,
- the *binary connectives* $\wedge$ and $\vee$,
- parentheses and brackets of various sizes.

The set of sentences of $\mathcal{L}$ is given by the following rules:

- Every propositional variable is a sentence. We call these the *atomic sentences* of $\mathcal{L}$.
- If ϕ is a sentence, then so is $\neg\phi$.
- If ϕ and ψ are sentences, then so are $(\phi \wedge \psi)$ and $(\phi \vee \psi)$.

As usual, we will take '$\neg$' to mean 'not', take '$\wedge$' to mean 'and', and take '$\vee$' to mean 'or'. That is, we take the truth-and-falsity conditions for these connectives to be given as follows:

- The sentence $\neg\phi$ is true if and only if ϕ is false.
- The sentence $\neg\phi$ is false if and only if ϕ is true.
- The sentence $(\phi \wedge \psi)$ is true if and only if ϕ is true and ψ is true.
- The sentence $(\phi \wedge \psi)$ is false if and only if at least one of ϕ and ψ is false.
- The sentence $(\phi \vee \psi)$ is true if and only if at least one of ϕ and ψ is true.
- The sentence $(\phi \vee \psi)$ is false if and only if ϕ is false and ψ is false.

With these things out of the way, let's now turn to seeing what the Standard Story tells us about the logic of $\mathcal{L}$.

Logical Consequence

The first – and as it turns out, most powerful – of the assumptions in the Standard Story is about the definition of logical consequence. As we stated it, logical consequence is defined in terms of *im*possibility: Δ is a logical consequence of Γ if and only if it is impossible for every sentence in Δ to be true while every sentence in Γ is not true.

There are two things worth remarking about this definition. The first is that it tells us what the *subject matter* of logic is: logic is about logical consequence, which is defined in the first of the standard assumptions. This view of logic is, somewhat surprisingly, both fairly new and fairly old. John Etchemendy has remarked on this in the following passage:

> Throughout much of this century, the predominant conception of logic was one inherited from Frege and Russell, a conception according to which the primary subject of logic, like the primary subject of arithmetic or geometry, was a particular body of truths: logical truths in the former case, arithmetical or geometric in the latter [...] This conception of logic now strikes us as rather odd, indeed as something of an anomaly in the history of logic. We no longer view logic as having a body of truths, the logical truths, as its principal concern; we do not, in this respect, think of it as parallel to other mathematical disciplines. If anything, we think of the consequence relation itself as the primary subject of logic, and view logical truth as simply the degenerate instance of this relation: logical truths are those that follow from any set of assumptions whatsoever, or alternatively, from no assumptions at all. (Etchemendy, 1988, 74)

The second is that, while this definition is not the only option on the table once we've settled on logical consequence as the subject matter of logic, it is nonetheless fairly clear that truth transmission is one of the *core components* of what logic is about. But the definition leaves open the issue of saying exactly what "possibilities" are. It seems like this could be a subtle matter. It's reasonable, for example, to wonder whether possibilities – whatever they are – need to be *complete*, that is, whether they need to "describe", in some sense, a complete world, where all the facts are settled. Another question to answer is whether possibilities need to be *consistent*.

If we stick to the Standard Story, then these questions are already answered for us. The completeness assumption tells us that logical possibilities are complete. The consistency assumption tells us that logical possibilities are consistent. And, since the Standard Story doesn't force us to make any other assumptions, a natural first stab at a definition of logical possibility is this:

Attempted Definition: A *logical possibility* is any way of labelling some sentences in our language as true and labelling all (and only!) the others as false.

This is almost correct, but misses a crucial detail that is exemplified by the following observation: given what '$\wedge$' means, $\phi \wedge \psi$ can be true if and only if both ϕ and ψ are true. But the attempted definition doesn't require this, and so allows a "logical possibility" in which we label $\phi \wedge \psi$ as true even though we've also labelled ϕ as false. To rule out this sort of deranged "possibility" we adopt the following definition instead:

Definition: A *logical possibility* is a way of labelling some sentences in our language as true while labelling all (and only!) the others as false such that

(a) The sentence $\neg\phi$ is labelled true if and only if ϕ is labelled false.
(b) The sentence $\neg\phi$ is labelled false if and only if ϕ is labelled true.
(c) The sentence $(\phi \wedge \psi)$ is labelled true if and only if ϕ is labelled true and ψ is labelled true.
(d) The sentence $(\phi \wedge \psi)$ is labelled false if and only if at least one of ϕ and ψ is labelled false.
(e) The sentence $(\phi \vee \psi)$ is labelled true if and only if at least one of ϕ and ψ is labelled true.
(f) The sentence $(\phi \vee \psi)$ is labelled false if and only if ϕ is labelled false and ψ is labelled false.

Formal Semantics

This definition of 'logical possibility' suggests a natural way to build a mathematically precise formal semantics for $\mathcal{L}$.[1] To begin, say that a *model* of $\mathcal{L}$ is simply a function from $\mathcal{L}$ to the set $\{\top, \bot\}$ ('$\top$' for true and '$\bot$' for false) that obeys a few assumptions. Models will play the role of possibilities in our formal semantics. Taking models to be functions forces them to behave in accordance with the first part of our definition of 'logical possibility': a model (logical possibility) is an assignment of *exactly one* of the semantic statuses $\top$ (true) and $\bot$ (false) to every sentence.

[1] Eagle-eyed readers may also note that this definition is quite redundant. That's intentional, and will be useful later on.

To capture conditions (a)–(f), on the other hand, we require that models meet their obvious counterparts. Altogether, then, we define what a model is as follows.

Definition: A function $v : \mathcal{L} \to \{\top, \bot\}$ is a *model* if and only if it meets the following assumptions:

- $v(\neg\phi) = \top$ if and only if $v(\phi) = \bot$.
- $v(\neg\phi) = \bot$ if and only if $v(\phi) = \top$.
- $v(\phi \wedge \psi) = \top$ if and only if $v(\phi) = \top$ and $v(\psi) = \top$.
- $v(\phi \wedge \psi) = \bot$ if and only if $v(\phi) = \bot$ or $v(\psi) = \bot$.
- $v(\phi \vee \psi) = \top$ if and only if $v(\phi) = \top$ or $v(\psi) = \top$.
- $v(\phi \vee \psi) = \bot$ if and only if $v(\phi) = \bot$ and $v(\psi) = \bot$.

We can now give a mathematically precise account of formal consequence for $\mathcal{L}$ as follows: Δ *follows semantically* from Γ if and only if there is no model v such that $v(\phi) = \top$ for every $\phi \in \Gamma$ but for every $\psi \in \Delta$, $v(\psi) \neq \top$. When this happens, we write $\Gamma \vDash \Delta$. Since this is the account of consequence we get when we flesh out the standard assumptions, we call this classical logic (for $\mathcal{L}$).

A Comparison

In Chapter 1, we said that a model of a logic would be an 8-tuple $\langle W, \mathcal{R}, \mathcal{D}, \mathcal{V}, \mathcal{K}, \mathcal{C}, \delta, \mathfrak{v}\rangle$. We can understand the "models" we just presented as models in this sense by understanding the various elements as follows:

- W is a singleton set $\{\bullet\}$.
- $\mathcal{R}$ is arbitrary.
- $\mathcal{D}$ is arbitrary.
- $\mathcal{V} = \{\top, \bot\}$.
- $\mathcal{K} = \{\{\top\}, \{\bot\}\}$.
- $\mathcal{C}$ is the collection of functions given by the following '*truth-tables*':

v	$\rho_\neg(v)$
$\top$	$\bot$
$\bot$	$\top$

v_1	v_2	$\rho_\wedge(v_1, v_2)$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\bot$
$\top$	$\bot$	$\bot$
$\bot$	$\bot$	$\bot$

v_1	v_2	$\rho_\vee(v_1, v_2)$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\top$
$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\bot$

- Since we have no terms, δ can be the empty function.
- Finally, $\mathfrak{v}$ is an assignment of either $\top$ or $\bot$ to each propositional variable.

Examining classical logic

So we now have on our hands a relation of logical consequence. Like any other relation of logical consequence (see the discussion in Chapter 1), we can examine both its *structural features* and its *operational features*. Recall that in Chapter 1, we singled out the following structural rules as common to many of the logics we would study:

Reflexivity

$$\frac{}{\phi \Rightarrow \phi}$$

Weakening

$$\text{WL}\ \frac{\Gamma \Rightarrow \Delta}{\phi, \Gamma \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \phi}\ \text{WR}$$

Exchange

$$\text{EL}\ \frac{\Gamma, \phi, \psi, \Delta \Rightarrow \Pi}{\Gamma, \psi, \phi, \Delta \Rightarrow \Pi} \qquad \frac{\Gamma \Rightarrow \Delta, \phi, \psi, \Pi}{\Gamma \Rightarrow \Delta, \psi, \phi, \Pi}\ \text{ER}$$

Contraction

$$\text{CL}\ \frac{\phi, \phi, \Gamma \Rightarrow \Delta}{\phi, \Gamma \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \phi, \phi}{\Gamma \Rightarrow \Delta, \phi}\ \text{CR}$$

Cut

$$\frac{\Gamma \Rightarrow \Delta, \phi \qquad \phi, \Pi \Rightarrow \Sigma}{\Gamma, \Pi \Rightarrow \Delta, \Sigma}$$

Let's think now about which of these we'll accept if we read '$\Rightarrow$' as '$\vDash$'. As it turns out, there's not a lot of *logic* to think about when we do so. There's just a little bit of *set theory*. And the reason for this is simple: our definition of $\vDash$ ignores all but the set-like structure of Γ and Δ. That is, even if Γ and Δ were (somehow and for some reason) given to us with rich internal structure, it wouldn't matter. As we've defined '$\vDash$', all that matters is whether there is a model v that meets the following two conditions:

- For each $\phi \in \Gamma$, ϕ is true in v.
- For each $\psi \in \Delta$, ψ is false in v.

If there *is* such a model, Δ does *not* follow from Γ. If there is *not* such a model, Δ *does* follow from Γ. So, in particular, all that matters is what's going on with the *elements* of Δ and Γ. So we may as well just directly take Γ and Δ to be sets of sentences.

Interpretively, this requires two tweaks:

- When we write 'Γ, Δ' what we mean is '$\Gamma \cup \Delta$'.
- When we write 'Γ, ϕ', what we mean is '$\Gamma \cup \{\phi\}$'.

Given this interpretation, essentially all of the structural rules just listed are immediately validated. An example: the set $\Gamma \cup \{\phi\} \cup \{\phi\}$ is, for all sets Γ and sentences ϕ, exactly the same set as the set $\Gamma \cup \{\phi\}$. This validates the CL rule. Almost all the other rules are validated by similarly trivial arguments. The only one that needs a bit more is Cut. Here's how to see that it too is valid:

Proof: Suppose $\Gamma \vDash \Delta, \phi$ and $\phi, \Pi \vDash \Sigma$. Let v be a model that makes true everything in Γ and everything in Π. Then, since v is a model, v either makes ϕ true or makes ϕ false.

Suppose v makes ϕ true. Then since v also makes everything in Π true and Σ follows from $\{\phi\} \cup \Pi$, we can conclude that v must make something in Σ true. Thus, v makes something in Σ or something in Δ true.

Now suppose v makes ϕ false. Then since v makes everything in Γ true and $\Delta \cup \{\phi\}$ follows from Γ, we can conclude that v either makes something in Δ true or makes ϕ true. But v doesn't make ϕ true, because v makes ϕ false. So v makes something in Δ true, and thus either makes something in Σ or something in Δ true.

So, either way, the assumption that v makes true everything in Γ and everything in Π led us to the conclusion that v makes true something in Δ or something in Σ. So no model makes true everything in Γ and everything in Π while failing to make true anything in either Δ or Σ. So $\Gamma, \Pi \vDash \Delta, \Sigma$.

□

Now we turn to the operational rules. One can (though we leave it to the reader to do so) verify that all the operational rules for $\wedge$, $\vee$ and $\neg$ that were proposed in Chapter 1 remain valid when we interpret '$\Rightarrow$' as '$\vDash$'. But it turns out that, for this chapter, our lives will be much easier if we split our treatment of negated formulas into pieces. Thus, we will (only in this chapter!) deal with the following, much larger collection of operational rules:

Axioms

$$\frac{}{\phi \Rightarrow \phi}\,\text{id} \qquad \frac{}{\phi, \neg\phi \Rightarrow}\,\text{con} \qquad \frac{}{\Rightarrow \phi, \neg\phi}\,\text{com}$$

Conjunction Rules

$$\text{L}(\wedge)\,\frac{\Gamma, \phi \Rightarrow \Delta}{\Gamma, \phi \wedge \psi \Rightarrow \Delta} \qquad \frac{\Gamma, \psi \Rightarrow \Delta}{\Gamma, \phi \wedge \psi \Rightarrow \Delta}\,\text{L}(\wedge)$$

$$\frac{\Gamma \Rightarrow \phi, \Delta \qquad \Gamma \Rightarrow \psi, \Delta}{\Gamma \Rightarrow \phi \wedge \psi, \Delta}\,\text{R}(\wedge)$$

$$\frac{\Gamma, \neg\phi \Rightarrow \Delta \qquad \Gamma, \neg\psi \Rightarrow \Delta}{\Gamma, \neg(\phi \wedge \psi) \Rightarrow \Delta}\,\text{L}(\neg\wedge)$$

$$\text{R}(\neg\wedge)\,\frac{\Gamma \Rightarrow \neg\phi, \Delta}{\Gamma \Rightarrow \neg(\phi \wedge \psi), \Delta} \qquad \frac{\Gamma \Rightarrow \neg\psi, \Delta}{\Gamma \Rightarrow \neg(\phi \wedge \psi), \Delta}\,\text{R}(\neg\wedge)$$

Disjunction Rules

$$\text{R}(\vee)\,\frac{\Gamma \Rightarrow \phi, \Delta}{\Gamma \Rightarrow \phi \vee \psi, \Delta} \qquad \frac{\Gamma \Rightarrow \psi, \Delta}{\Gamma \Rightarrow \phi \vee \psi, \Delta}\,\text{R}(\vee)$$

$$\frac{\Gamma, \phi \Rightarrow \Delta \qquad \Gamma, \psi \Rightarrow \Delta}{\Gamma, \phi \vee \psi \Rightarrow \Delta}\,\text{L}(\vee)$$

$$\frac{\Gamma \Rightarrow \neg\phi, \Delta \qquad \Gamma \Rightarrow \neg\psi, \Delta}{\Gamma \Rightarrow \neg(\phi \vee \psi), \Delta}\,\text{R}(\neg\vee)$$

$$\text{L}(\neg\vee)\,\frac{\Gamma, \neg\phi \Rightarrow \Delta}{\Gamma, \neg(\phi \vee \psi) \Rightarrow \Delta} \qquad \frac{\Gamma, \neg\psi \Rightarrow \Delta}{\Gamma, \neg(\phi \vee \psi) \Rightarrow \Delta}\,\text{L}(\neg\vee)$$

Negation Rules

$$\neg I_L\,\frac{\Gamma \Rightarrow \phi, \Delta}{\Gamma, \neg\phi \Rightarrow \Delta} \qquad \frac{\Gamma, \phi \Rightarrow \Delta}{\Gamma \Rightarrow \neg\phi, \Delta}\,\neg I_R$$

Again, one can check rule-by-rule that if we take '$\Rightarrow$' to mean '$\vDash$', then all of these are in fact consequences of the given semantic theory. For example, take the $R(\neg\vee)$ rule. Here is an argument establishing its acceptability on the semantic theory just given:

Proof Suppose $\Gamma \vDash \neg\phi, \Delta$ and $\Gamma \vDash \neg\psi, \Delta$. Let v be a model and suppose that v makes true every $\phi \in \Gamma$. Then since $\Gamma \vDash \neg\phi, \Delta$ we conclude that v must make true either $\neg\phi$ or something in Δ. And since $\Gamma \vDash \neg\psi, \Delta$ we conclude that v must make true either $\neg\psi$ or something in Δ.

If v makes true something in Δ, then v makes true something in $\{\neg(\phi \vee \psi)\} \cup \Delta$. On the other hand, if v does not make true something in Δ, then v must make true $\neg\phi$ and make true $\neg\psi$. So v makes false both ϕ and ψ. So it makes false $\phi \vee \psi$. Thus it makes true $\neg(\phi \vee \psi)$. So it makes true something in $\{\neg(\phi \vee \psi)\} \cup \Delta$. So $\Gamma \vDash \neg(\phi \vee \psi), \Delta$. □

Alternatives

Between the semantics and the proof theory, we now have a fairly good grasp on classical logic – at least as far as $\mathcal{L}$ is concerned. The question to ask at this point, then, is whether classical logic is the logic we want. To examine this, we'll focus on something that has been conspicuously absent from our discussion so far: the conditional.

In English, of course, the conditional is indicated by the phrase 'if… then…'. $\mathcal{L}$ contains no obvious analogue of this connective. We can, however, say a little bit about how such a connective would have to behave, and then turn to examining what the Standard Story says such a connective will be like.

How a Conditional Has to Behave

We'll now focus on a new language $\mathcal{L}^+$ that extends $\mathcal{L}$ by including an additional connective – which we will write '$\rightarrow$' – that will be our formal counterpart to the 'if… then…' conditional in English. Technically, this means that $\mathcal{L}^+$ differs from $\mathcal{L}$ both in its vocabulary and in its set of grammatical rules, but it's safe to leave the technicalities aside here, so we will.

At any rate, for $\rightarrow$ to play the 'if… then…' role well, it will have to satisfy the following two conditions:

1. Given sentences ϕ and ψ, whenever ϕ is true and ψ is false, $\phi \rightarrow \psi$ must also be false.
2. Given sentences ϕ and ψ and a set of sentences Γ, whenever ψ follows from Γ taken together with ϕ, $\phi \rightarrow \psi$ must follow from Γ alone.

It's intuitively clear that anything playing the 'if... then...' role will have to satisfy these conditions. What's surprising is that according to the Standard Story if $\rightarrow$ obeys these rules then for any two sentences ϕ and ψ, the sentence $\phi \rightarrow \psi$ will be equivalent to the sentence $\neg\phi \vee \psi$. Since this is a rather surprising conclusion, we'll pause to prove it:

Proof First note that ϕ is true iff $\neg\phi$ is false. Since models are functions, it follows that there is no model in which both ϕ and $\neg\phi$ are true. So there is no model in which ϕ and $\neg\phi$ are true but ψ is false. So by the definition of logical consequence, ψ follows from $\neg\phi$ taken together with ϕ. Thus by condition 2, $\phi \rightarrow \psi$ follows from $\neg\phi$. So, again by the definition of logical consequence, whenever $\neg\phi$ is true, $\phi \rightarrow \psi$ is true.

On the other hand, it's obvious that whenever ψ and ϕ are both true, ψ is true. So by the definition of logical consequence, ψ follows from ψ taken together with ϕ. Thus, again by condition 2, $\phi \rightarrow \psi$ follows from ψ. So whenever ψ is true, $\phi \rightarrow \psi$ is true.

Altogether, whenever $\neg\phi$ is true or ψ is true, $\phi \rightarrow \psi$ is true, and thus whenever $\neg\phi \vee \psi$ is true, $\phi \rightarrow \psi$ is true.

On the other hand, the only way for $\neg\phi \vee \psi$ to be false is for $\neg\phi$ to be false and ψ to be false. But then, by the truth-conditions for the negation, ϕ is true and ψ is false. But by condition 1, $\phi \rightarrow \psi$ is also false when this happens.

Altogether, whenever $\neg\phi \vee \psi$ is true, so is $\phi \rightarrow \psi$ and whenever $\neg\phi \vee \psi$ is false, so is $\phi \rightarrow \psi$, so the two are equivalent. □

It's common to call a connective $*$ a *material conditional* when $\phi * \psi$ is, for all sentences ϕ and ψ, equivalent to $\neg\phi \vee \psi$. So what we've just established is that it follows from the Standard Story that only material conditionals can play the role of the conditional.

Many logicians and philosophers have thought that this is an absurd conclusion. Here's one such complaint:

> Let us suppose that Roy Dyckhoff has claimed that John Slaney was in Edinburgh on a certain day, and that Crispin Wright has denied it. Consider the following three propositions as they describe this situation:
>
> (1) If John was in Edinburgh, Roy was right.
>
> This is clearly true: that's what Roy claimed.

> (2) If Crispin was right, so was Roy.
>
> That is equally obviously false, given the logic of denial.
>
> (3) If John was in Edinburgh, Crispin was right.
>
> That too is false, for Crispin denied it. Let us use these propositions to construct an argument, taking as premises (1) together with the denial of (2), and as conclusion (3):
>
> If John was in Edinburgh, then Roy was right.
> It's not the case that if Crispin was right, so was Roy.
> Hence, if John was in Edinburgh, Crispin was right.
>
> Since (1) is true and (2) and (3) false, this argument, which takes the denial of (2) as its second premise, has true premises and a false conclusion. Hence it is invalid.
> Classically, however, the argument is valid. For the sequent
>
> $$\begin{array}{l} \quad P \to Q \\ \quad \neg(R \to Q) \\ \therefore P \to R \end{array}$$
>
> which formalises the argument classically... is (classically) valid. Hence, if the truth-conditions of 'if' were correctly captured by material implication, and the Classical Account of Validity were right, the argument would be valid. But it is not. So either treating 'if' as truth-functional, or the Classical Account of validity, is wrong. (Read, 2012, 23–24)

One has to hand it to the dissident: by all lights, the offset argument in the above passage *looks* to be invalid. On the other hand, it doesn't feel like the argument that established the materiality of the conditional went wrong anywhere. Thus, if we want to reject the equivalence of $\phi \to \psi$ and $\neg\phi \vee \psi$, we have to reject (at least) one of its hypotheses. But all we seemed to rely on were (a) the Standard Story and (b) the characterization we gave above of what it takes to play the role of a conditional. So, if we don't want to accept the materiality of the conditional, then we must either challenge the Standard Story or challenge this characterization of the conditional. We will focus our discussion on the former approach.

Gluts

Go back and have a look at the argument we gave on page 44 for the materiality of the conditional again. One thing that might strike you is the prominent role played by Consistency. If you're predisposed to rejecting the conclusion of this argument, then this casts Consistency in a pretty bad light. Moreover, we can, it seems, build explicit counterexamples to Consistency. Consider, for example, the following sentence:

⋆ The sentence preceded by a star on page 46 is false.

If this sentence is true, then it's also false, because that's what it says about itself. On the other hand, if this sentence is false, then it's true, because (again) that's just what it says! So either way, it ends up being both true and false. So it's a counterexample to consistency.

If you felt like that argument was too quick, you're in good company. This sentence – it's known as the Liar Sentence – has been subject to much scrutiny since ancient times. But the Liar is just one part of a much larger fight about consistency that has been brewing for quite some time.

Central to the debate about consistency is the question of the validity of the *law of noncontradiction*. Aristotle, it turns out, had a good bit to say on the subject. For example, Paula Gottlieb has usefully singled out three different versions of the law of noncontradiction discussed by Aristotle in *Metaphysics IV* (Gottlieb, 2015):

- It is impossible for the same thing to belong and not to belong at the same time to the same thing and in the same respect.
- It is impossible to hold (suppose) the same thing to be and not to be.
- Opposite assertions cannot be true at the same time.

Whether these three are really *the same* in the relevant sense is a thorny issue that we won't get into. Instead, it's worth thinking through how we might formalize these in $\mathcal{L}$. There are, it seems to us, three good options about what noncontradiction might amount to. It might amount to

(a) $\phi \wedge \neg\phi$ never being true,
(b) $\neg(\phi \wedge \neg\phi)$ being logically true, or
(c) the claim that every argument of the following form is valid:

$$\begin{array}{ll} & \phi \\ & \neg\phi \\ \therefore & \Gamma \end{array}$$

The first captures the intuition that $\phi \wedge \neg\phi$ is "impossible". The second captures the intuition that $\phi \wedge \neg\phi$ is always *false*. The third cashes out an intuition in the vicinity of these in terms of our definition of logical consequence: if it's impossible for ϕ and $\neg\phi$ to be true, then, for any ψ, it must be impossible for ϕ and $\neg\phi$ to be true while everything in Γ is false. So, by our definition of logical consequence, the argument from ϕ and $\neg\phi$ to Γ is valid.

We'll return to these in a bit. For now, the thing to focus on is this: both the Liar Sentence and the role Consistency plays in our proof that the conditional must be material give us good reasons to at least *explore* what are called *para*consistent logics – logics that reject the consistency assumption. To get a feel for how the paraconsistent approach to logic goes, we will explore one such logic in some detail. The logic we present is, in many ways, the most straightforward paraconsistent logic possible. It's known as the *logic of paradox*, or just LP, and it was introduced by Graham Priest (1979).

Paraconsistency: Semantics

Let's begin by restricting our attention to $\mathcal{L}$ again. Semantically, what distinguishes LP from classical logic is this: instead of models being functions from $\mathcal{L}$ to the set $\{\top, \bot\}$, models will be functions from $\mathcal{L}$ to $\{\top, \bot, \mathrm{b}\}$, where the 'b' semantic status is taken to be both true and false. Thus b is the semantic value taken by all glutty sentences. Thus, in an LP-model, no sentence is neither true nor false, but some sentences are both true and false.

In terms of the model-theoretic characterization of logics from Chapter 1, we have the following:

- W, $\mathcal{R}$, $\mathcal{D}$, and δ are as before.
- $\mathcal{V} = \{\top, \bot, \mathrm{b}\}$.
- $\mathcal{K} = \{\{\top, \mathrm{b}\}, \{\bot, \mathrm{b}\}\}$.
- $\mathcal{C}$ is the collection of functions given by the following *truth-tables*:

v	$\rho_{\neg}^{LP}(v)$
$\top$	$\bot$
$\bot$	$\top$
b	b

v_1	v_2	$\rho_{\wedge}^{LP}(v_1, v_2)$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\bot$
b	$\top$	b
$\top$	$\bot$	$\bot$
$\bot$	$\bot$	$\bot$
b	$\bot$	$\bot$
$\top$	b	b
$\bot$	b	$\bot$
b	b	b

v_1	v_2	$\rho_{\vee}^{LP}(v_1, v_2)$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\top$
b	$\top$	$\top$
$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\bot$
b	$\bot$	b
$\top$	b	b
$\bot$	b	b
b	b	b

- $\mathfrak{v}$ is again determined recursively, this time beginning with an assignment of $\top$, $\bot$ or b to each propositional variable.

The truth-tables above might appear somewhat mysterious. They're best explained by first examining $\mathcal{K}$. Notice that b occurs in both the set of designated values and the set of antidesignated values. So it counts both as a way to be true and as a way to be false. With this in mind, we can understand the functions as a way of spelling out the definition of logical possibilities given on p. 38.

Paraconsistency: Proof Theory

We can characterize LP proof-theoretically using a simple modification of our proof theory for classical logic. All we have to do is omit the *con* rule from our list of axioms and the result is a sound and complete proof theory for LP.[2] We won't bother to prove this here, though a proof of a similar result can be found in Beall (2011) or in Priest (2008), among other places.

Examining LP

So now we have a different logic – a *rival* to classical logic – on our hands. One question worth looking into is just how different LP actually is. And

[2] A proof theory is *sound* with respect to a model theory iff, when $\Gamma \vdash \phi$ then $\Gamma \vDash \phi$. A proof theory is *complete* with respect to a model theory iff, when $\Gamma \vDash \phi$ then $\Gamma \vdash \phi$. See Chapter 3 for a discussion of these properties.

there's one sense in which the answer is that it's not different at all. Given a logic L, we say that a sentence ϕ is an *L-logical truth* when ϕ receives a designated value in every model. LP and classical logic, it turns out, have exactly the same logical truths:

Proof Suppose ϕ is not an LP-logical truth, and let v be an LP-model witnessing this; that is, so that $v(\phi) = \bot$. Define a function v' from propositional variables to $\{\top, \bot\}$ as follows:

$$v'(p) = \begin{cases} \top & \text{if } v(p) = \top \\ \bot & \text{otherwise} \end{cases}$$

So v' makes false each propositional variable that is at least false in v. We can then recursively extend v' to a classical model in the usual way. When we do so, it turns out (we leave it to the reader to check this) that, for each $\phi \in \mathcal{L}$, $v'(\phi) = \top$ iff $v(\phi) = \top$ and $v(\phi) = \bot$ otherwise.

Thus, since $v(\phi) = \bot$, $v'(\phi) = \bot$. But v' is a classical model. So if ϕ is not an LP-logical truth, then ϕ is also not a classical logical truth.

On the other hand, note that if ϕ is not a classical logical truth, then for some model v, $v(\phi) = \bot$. But every classical model *already is* an LP-model. Thus ϕ is not an LP-logical truth. □

This fact definitively nixes option (b) for how to translate the law of non-contradiction: it certainly isn't correct to formalize it as *just* saying that $\neg(\phi \wedge \neg\phi)$ is a logical truth. The reason is that LP is *explicitly* formulated to *allow* contradictions, yet nonetheless, since $\neg(\phi \wedge \neg\phi)$ is a classical logical truth, it's also an LP-logical truth. On the other hand, options (a) and (c) are still live matters: LP, unlike classical logic, doesn't rule out $\phi \wedge \neg\phi$ being true. It *does* agree with classical logic that this sentence is always false; it just rejects that falsity and untruth are tightly paired, and it thus allows that this sentence might, despite being false, also be true. And LP flat-out rejects that the argument from ϕ and $\neg\phi$ to Γ is (generally) valid.[3] Thus, despite the fact that LP and classical logic share all the same theorems it turns out that LP differs from classical logic in rather substantive ways.

There's a cautionary tale to be told here, though: when you throw out a few things in a logic, often other things you'd like to have kept end up going out with them. An example: in classical logic, one can validly infer

[3] Notice that for some Γs, LP is fine with this argument – e.g. if Γ is just ϕ or just $\neg\phi$.

ψ from $\phi \vee \psi$ and $\neg\phi$. This is because classical models are functions. Thus, any model that makes $\neg\phi$ true must make ϕ false. So if this same model makes $\phi \vee \psi$ true, then it must be because it makes ψ true.

But LP-models aren't functions; they're relations. So the fact that an LP-model v makes $\neg\phi$ true (that is, such that $v \models_1 \neg\phi$) does not rule out v making ϕ true as well – all that's required is that ϕ be related to both $\top$ and $\perp$. But this can happen while ψ is related only to $\perp$. Any such model is then a counterexample to disjunctive syllogism.

What this shows is something interesting: the paraconsistent route to avoiding the conclusion that all conditionals are material runs the risk of invalidating rather commonsensical inferences (like disjunctive syllogism, for example). This may seem a heavy cost to pay. There is, however, a bit of a silver lining, which comes, as is frequently the case in philosophical logic, in the form of a theorem, this one due to Jc Beall:

Theorem: If $\Gamma \vdash \Delta$ is a classically valid inference, then $\Gamma \vdash \Delta \cup \iota(\Gamma)$ is LP-valid, where we define $\iota(\Gamma)$ recursively as follows:

- If ϕ is an atomic sentence, then $\iota(\{\phi\}) = \{\phi \wedge \neg\phi\}$.
- For non-atomic sentences ϕ, we define ι as follows:
  - $\iota(\{\neg\phi\}) = \iota(\{\phi\})$
  - $\iota(\{\phi \wedge \psi\}) = \iota(\{\phi \vee \psi\}) = \iota(\phi) \cup \iota(\psi)$
- Finally, for sets of sentences, we define ι as follows:
  - $\iota(\Gamma) = \bigcup\limits_{\phi \in \Gamma} \iota(\phi)$.

$\iota(\Gamma)$ is known as the *inconsistency set* of Γ. Each element of $\iota(\Gamma)$ is of the form $\ulcorner\phi \wedge \neg\phi\urcorner$[4] for some atomic sentence ϕ that occurs as part of some member of Γ. And $\iota(\Gamma)$ contains all such "atomic contradictions".

We won't prove this theorem here, though we will remark that a fairly gentle proof of the result can be found in Beall (2011). What the theorem tells us is that, essentially, LP is no more than a *cautious version* of classical logic. Where classical logic endorses $\Gamma \vdash \Delta$ directly, LP proceeds a bit more

[4] Throughout the book we use single quotation marks for *mentioning* quoted expressions and double quotation marks when we *use* a quoted expression. We use corner quotes, as in this case, when part of a quoted expression that is mentioned is still used. Here it is the metavariable 'ϕ', which occurs in a mentioned phrase, but is bound from "outside" the quoted expression.

cautiously, endorsing instead the claim that either (a) $\Gamma \vdash \Delta$ holds *or* (b) something in Γ is a contradiction.

All of that is to say that there are pros and cons to adopting LP as our preferred logic. But what of the conditional, which motivated this detour in the first place? Here, it turns out, there is a problem: it can be shown (see, e.g. Beall et al., 2013) that LP cannot be equipped with a meaningful connective that can play the 'if… then…' role in the sense we characterized above. In short, the LP-semantics *rule out* our adding to $\mathcal{L}$ a connective that can play the 'if… then…' role. Of course, we can perfectly well add a *material conditional* to LP, and if we do so, we can "recover" all the classical theorems about this connective in the sense that we can "recover" any other classical theorem as given in the above theorem. But unless we're *also* willing to challenge the characterization of the conditional we gave above, LP is definitively not a way to avoid the conditional being material.

Gaps

Before we carry on, again go back and look at our argument for the materiality of the conditional. As we noted before, consistency plays a prominent role here. Less obvious is that *completeness* also plays a role. Here's how: what the argument shows, explicitly, is that if a connective $*$ can play the 'if… then…' role, then for any sentences ϕ and ψ, the sentence $\phi * \psi$ will be true (or false) when $\neg\phi \vee \psi$ is true (or false). But notice that unless we assume that there are no other options than truth and/or falsity for a sentence, this alone doesn't guarantee that $\phi * \psi$ and $\neg\phi \vee \psi$ are *equivalent*, since it doesn't settle the matter of what to do when one of them is neither true nor false.

So an alternative proposal for how to avoid the equivalence of $\phi \rightarrow \psi$ with $\neg\phi \vee \psi$ is to adopt a logic that rejects *completeness* (and so is a para*complete*, rather than a para*consistent* logic). As with paraconsistent approaches, there is a lot of philosophical history here.

The key to the philosophical content of the completeness assumption is to notice its connection to the law of the excluded middle: roughly the claim that every sentence is either true or false. And, as before, we will highlight three different ways to understand what this means. It could mean that

(a) $\phi \vee \neg\phi$ is always true,

(b) $\neg(\phi \vee \neg\phi)$ is logically false, or
(c) every argument of the following form is valid:

$$\begin{array}{l} \quad \Gamma \\ \therefore \ \phi \vee \neg\phi \end{array}$$

As before, these cash out different intuitions in the vicinity of excluded middle. The first captures the idea that either ϕ or $\neg\phi$ must always be true. The second captures the intuition that the simultaneous falsity of both ϕ and $\neg\phi$ is impossible. The third cashes things out in terms of logical consequence: if one of ϕ and $\neg\phi$ is always true, then for any set of sentences Γ, it is impossible for everything in Γ to be true without at least one thing in $\{\phi, \neg\phi\}$ being true.

We will examine which of these does the job of capturing the intuition behind excluded middle best as we go. However, before we turn to doing so, it's worthwhile reflecting on something we learned in our previous excursion into nonclassical logics: untruth and falsity can, in such settings, fall apart. So where previously we might have struggled to see a difference between these three formulations of excluded middle, we're now in position to mark these differences fairly well. This is a worthwhile exercise, so we won't complete it for you.

Paracompleteness: Semantics

As before, what we will be constructing is essentially the very simplest paracomplete logic possible. And, also as before, we begin by restricting our attention to $\mathcal{L}$. Semantically, what distinguished LP from classical logic is this: instead of models being functions from $\mathcal{L}$ to the set $\{\top, \bot\}$, models were functions from $\mathcal{L}$ to $\{\top, \bot, \mathrm{b}\}$, where the 'b' semantic status was taken to be both true and false.

The logic we are constructing now is similar in that it also takes there to be three truth-values. But this time we write this three element set as $\{\top, \bot, \mathrm{n}\}$, and take the 'n' semantic status to be neither true nor false.

In terms of the model-theoretic characterization of logics from Chapter 1, we have the following:

- W, $\mathcal{R}$, $\mathcal{D}$, and δ are as before.
- $\mathcal{V} = \{\top, \bot, \mathrm{n}\}$.
- $\mathcal{K} = \{\{\top\}, \{\bot\}\}$.

- $\mathcal{C}$ is the collection of functions given by the following *truth-tables*:

v	$\rho^{K3}_{\neg}(v)$
$\top$	$\bot$
$\bot$	$\top$
n	n

v_1	v_2	$\rho^{K3}_{\wedge}(v_1, v_2)$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\bot$
n	$\top$	n
$\top$	$\bot$	$\bot$
$\bot$	$\bot$	$\bot$
n	$\bot$	$\bot$
$\top$	n	n
$\bot$	n	$\bot$
n	n	n

v_1	v_2	$\rho^{K3}_{\vee}(v_1, v_2)$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\top$
n	$\top$	$\top$
$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\bot$
n	$\bot$	n
$\top$	n	$\top$
$\bot$	n	n
n	n	n

- $\mathfrak{v}$ is again determined recursively, this time beginning with an assignment of $\top$, $\bot$ or n to each propositional variable.

The truth-tables above are again ways of spelling out the definition of logical possibilities given on p. 38, this time with the understanding that n is neither a way of being true nor a way of being false.

The resulting logic is known as K3. The ‘K’ comes from Stephen **K**leene, who is credited with introducing this logic (Kleene, 1952).

Paracompleteness: Proof Theory

We characterized LP proof-theoretically by dropping one axiom from our proof-theoretic characterization of classical logic. We can, it turns out, also characterize K3 proof-theoretically by dropping one axiom. For our paraconsistent theory, we dropped the consistency assumption and, with it, the *con* axiom. For the paracomplete theory we drop the completeness assumption and, with it, the *com* axiom. That this gives a sound and complete proof theory for K3 can be proved by mildly modifying the proofs in Priest (2008) or (more directly) in the supplement to Beall and Logan (2017).

Examining K3

The pro/con calculus for K3 is essentially dual to that for LP. LP, we saw, had exactly the same theorems as classical logic. K3, it turns out, has no logical truths. None at all. This might seem somewhat surprising, but it can be seen to be true with minimal effort. First, notice that the function assigning each propositional variable to n extends to a K3-model in the usual way. Call this model 'E'. Next, notice that the truth-tables given above guarantee that in this model every sentence (atomic or otherwise) will be assigned the semantic status n. But this just means that no sentence is true in this model (and no sentence is false in this model either). Thus, for every sentence, there is at least this one model in which it isn't true. So no sentence is a logical truth.

We said above that one way to judge whether excluded middle held was by looking at whether $\phi \vee \neg\phi$ was a logical truth, and that another way was by looking at whether $\neg(\phi \vee \neg\phi)$ was a logical falsehood. This theorem tells us that both of these tests are effective: excluded middle clearly fails in K3, and, indeed, there are models in which $\phi \vee \neg\phi$ is not true (in particular, in E), and there are models in which $\neg(\phi \vee \neg\phi)$ is not false (again, E works).

On the other hand, where LP *rejects* many common-sense inference rules, these tend to be perfectly fine in K3. Thus, for example, *disjunctive syllogism* – the argument from $\phi \vee \psi$ and $\neg\phi$ to ψ – which is *in*valid in LP, is valid in K3.

To see this, suppose v is a K3 model in which $\phi \vee \psi$ and $\neg\phi$ are true. Then since $\phi \vee \psi$ is true, either $v(\phi) = \top$ or $v(\psi) = \top$. But $v(\phi) = \bot$, because $v(\neg\phi) = \top$. Thus, since v is a function, $v(\phi) \neq \top$. So $v(\psi) = \top$.

A final bit of duality: above we covered Jc Beall's explanation of how to see LP as a *cautious* version of classical logic in the sense that whenever $\Gamma \vdash \Delta$ is classically valid, $\Gamma \vdash \Delta \cup \iota(\Gamma)$ is LP-valid. Beall has also shown a similar result holds for K3: whenever $\Gamma \vdash \Delta$ is classically valid, $\epsilon(\Delta) \cup X \vdash \Delta$ is K3-valid. The duality here is quite deep: the definition of ϵ is again a sort of dual to the definition of ι.

Definition: We define $\epsilon(\Gamma)$ recursively as follows:

- If ϕ is an atomic sentence, then $\epsilon(\{\phi\}) = \{\phi \vee \neg\phi\}$.
- $\epsilon(\{\neg\phi\}) = \epsilon(\{\phi\})$
- $\epsilon(\{\phi \wedge \psi\}) = \epsilon(\{\phi \vee \psi\}) = \epsilon(\phi) \cup \epsilon(\psi)$
- $\epsilon(\Gamma) = \bigcup_{\phi \in \Gamma} \epsilon(\phi)$

So where classical logic endorses $\Gamma \vdash \Delta$ directly, K3 endorses the weaker conclusion that Δ follows from Γ *assuming* that everything in Γ is either true or false.

Table 2.1 *What should a three-valued truth-table for the conditional look like?*

ϕ	ψ	$\phi \to \psi$
$\top$	$\top$	$\top$
$\top$	n	?
$\top$	$\bot$	$\bot$
n	$\top$	$\top$
n	n	?
n	$\bot$	?
$\bot$	$\top$	$\top$
$\bot$	n	$\top$
$\bot$	$\bot$	$\top$

Altogether, the pros and cons of K3 are dual to those of LP, as we said. But what of the conditional? The argument we gave for the materiality of the conditional still holds in K3, though it's somewhat weakened. What it tells us is *only* what happens to $\phi \to \psi$ when either (a) ϕ is false, (b) ψ is true or (c) ϕ is true and ψ is false. This leaves us with quite a bit of freedom. To see this, consider Table 2.1, in which we have filled in everything so far settled, and left question marks in the open options (n marks when a sentence is neither true nor false). There are three open options here. For each, there are three options: $\top$, $\bot$ or n. So, from a technical perspective, there are 27 different potential conditionals to look through. We won't take a tour of all the options, as that would be rather tedious. But it's worth remarking that many of these potential conditionals have been explored, and the general theme has been this: if you are ready to accept the other baggage that comes with K3, then one or another of them might suit you. If not – and, in particular, if losing *all* your theorems is a problem – then it's likely that none of them will.

Gaps and Gluts

Of course, one doesn't *have to be* bothered by any of the oddities of LP or K3. And, if one isn't, then why not go whole-hog and accept that there might not only be truth-value gaps but also truth-value gluts?

One can do this. Semantically, the resulting logic is characterized as follows:

- W, $\mathcal{R}$, $\mathcal{D}$, and δ are as before.
- $\mathcal{V} = \{\top, \bot, \mathrm{b}, \mathrm{n}\}$.
- $\mathcal{K} = \{\{\top, \mathrm{b}\}, \{\bot, \mathrm{b}\}\}$.
- $\mathcal{C}$ is the collection of truth-functions one expects given the interpretations of the semantic values. We won't reproduce them here since they are somewhat large (16 rows for each of the binary connectives).
- $\mathfrak{v}$ is again determined recursively, this time beginning with an assignment of $\top$, $\bot$, b or n to each propositional variable.

Proof-theoretically, first-degree entailment (FDE) is characterized by dropping both the *com* axiom and the *con* axiom. The logic this characterizes is known as the logic of first-degree entailments, or simply FDE. A version of Beall's "classical collapse" holds for FDE as well: if $\Gamma \vdash \Delta$ is classically valid, then $\varepsilon(\Delta) \cup X \vdash \Delta \cup \iota(\Gamma)$ is valid in FDE. This was proved by Beall (2013).

FDE was introduced by Belnap (1977) and is, as it turns out, quite an interesting system, about which there is much more to say. But from the perspective being given here, there's nothing terribly important about it, so we'll leave it at that for now.

What About the Definition of Logical Consequence?

So far we've looked at rivals to classical logic that are generated by dropping the consistency assumption or dropping the completeness assumption. Given what we've seen so far, there's reason to worry that neither approach is terribly promising. One should be cautious in drawing this conclusion – we've examined what is, in essence, the most simplistic and naïve version of each of these approaches. It should hopefully be clear that much more sophisticated options are available. This is especially true once the full machinery of the characterization of logic given in Chapter 1 is brought to bear.

But of course, the Standard Story has a further assumption as well – the definition of logical consequence. And it is, as we previously noted, a very *powerful* assumption. In Chapter 1, we saw one alternative notion worth considering: D^--logical consequence. It's a worthwhile exercise for the reader to think through what this amounts to in each of the above options. We won't have more to say about this matter here, though, since we'll be covering several options later in the book – in particular in Chapters 7 and 9.

Instead, we'll now turn to looking at *first-order* rivals to classical logic. As it turns out, there are also uniquely *higher-order* ways to motivate logical rivalry as well (see, for example, the discussion of Henkin models in Chapter 4), but we will leave those aside for now as well.

First-Order Rivals

As before, being very explicit and cautious pays dividends, so our first step will be to specify our first-order formal syntax. We will call the new language we are discussing $\mathcal{F}$, for **f**irst order. Its vocabulary consists of a countably infinite set of *individual constants* (we will use $a, b, \ldots, v, a_1, b_1$, etc. for these); a countably infinite set of *individual variables* (we will use w, x, y, z, w_1, etc. for these); the *identity symbol* $=$; for each n a countably infinite set of n-ary predicates (we will use P^n, Q^n, etc. for these); the connectives $\wedge$, $\vee$ and $\neg$; and, finally, the quantifiers $\forall$ and $\exists$.

We define the class of sentences in several stages. First, we define the class of well-formed formulas (wffs) recursively as follows:

- If τ_1 and τ_2 are either individual constants or individual variables, then $\tau_1 = \tau_2$ is an atomic wff.
- If Π^n is an n-ary predicate and $\tau_1, \tau_2, \ldots, \tau_n$ are all either individual constants or individual variables, then $\Pi^n\tau_1\tau_2\ldots\tau_n$ is an atomic wff.
- An atomic wff is a wff.
- If ϕ is a wff, then so is $\neg\phi$.
- If ϕ and ψ are wffs, then so are $(\phi \wedge \psi)$ and $(\phi \vee \psi)$.
- If ϕ is a wff and ν is a variable, then $\forall\nu\phi$ and $\exists\nu\phi$ are wffs.

We next define when an occurrence of a variable in a wff is a free occurrence of that variable.

- If ϕ is an atomic wff and ν is a variable, then each occurrence of ν in ϕ is a free occurrence.
- An occurrence of ν in $\neg\phi$ is a free occurrence just if it is a free occurrence of ν in ϕ.

- An occurrence of ν in $(\phi \wedge \psi)$ or in $(\phi \vee \psi)$ is a free occurrence just if it is a free occurrence of ν in ϕ or a free occurrence of ν in ψ.
- No occurrence of ν is free in $\forall\nu\phi$ or in $\exists\nu\phi$.

Finally, we define a sentence to be a wff in which no variable occurs freely.

Semantics

We interpret the language of first-order logic in first-order models. As before, we will first explain these in a free-standing way, then we will relate this explanation to the description of model-theoretic semantics from Chapter 1.

As we will understand it, a model has four pieces:

- a non-empty set D called the *domain* of the model;
- a function δ mapping the set of individual constants into D;
- a function $\mathfrak{v}$ mapping each n-ary predicate Π^n other than identity to a function from D^n to $\{\top, \bot\}$.

We confess that this definition of an interpretation is (at least mildly) idiosyncratic: usually the functions δ and $\mathfrak{v}$ are bundled together into a single "interpretation function". It is also the case that the function $\mathfrak{v}$ is usually defined somewhat differently. But for a variety of reasons, we find it more intuitive and explanatory to define interpretations in the way we have.

If α is a name, then $\delta(\alpha)$ is called the *denotation* of α. If Π^n is an n-ary predicate, then the *extension* of Π^n, written $\mathcal{E}^+(\Pi^n)$, and the *antiextension* of Π^n, written $\mathcal{E}^-(\Pi^n)$, are defined as follows:

$$\mathcal{E}^+(\Pi^n) = \{\langle d_1, \ldots, d_n \rangle \in D^n : \mathfrak{v}(\Pi^n)(\langle d_1, \ldots, d_n \rangle) = \top\}$$

$$\mathcal{E}^-(\Pi^n) = \{\langle d_1, \ldots, d_n \rangle \in D^n : \mathfrak{v}(\Pi^n)(\langle d_1, \ldots, d_n \rangle) = \bot\}$$

Finally, if M is a model, a *variable assignment for M* is a function va mapping the set of variables to the domain, D. If va is a variable assignment, ν is a variable, and $d \in D$, then by va_d^ν we mean the variable assignment defined as follows:

$$\text{va}_d^\nu(x) = \begin{cases} \text{va}(x) & \text{if } x \neq \nu \\ d & \text{if } x = \nu \end{cases}$$

Intuitively, va_d^ν is the variable assignment that is just like va except that it sends ν to d rather than to wherever va sent it.

If M is a model and va is a variable assignment, then we define a function ε from the set of individual constants and individual variables to D as follows:

$$\varepsilon(x) = \begin{cases} \text{va}(x) & \text{if } x \text{ is an individual variable} \\ \delta(x) & \text{if } x \text{ is an individual constant} \end{cases}$$

Intuitively, ε is just what we get when we "glue together" the variable assignment va and the denotation function δ.

Together, a model M and a variable assignment va give us enough information to recursively assign a semantic value $\mathfrak{v}_{\text{va}}^M(\phi)$ to each wff ϕ.

- $\mathfrak{v}_{\text{va}}^M(\tau_1 = \tau_2) = \begin{cases} \top & \text{if } \varepsilon(\tau_1) = \varepsilon(\tau_2) \\ \bot & \text{otherwise} \end{cases}$
- $\mathfrak{v}_{\text{va}}^M(\Pi^n\tau_1 \ldots \tau_n) = \mathfrak{v}(\Pi^n)(\langle \varepsilon(\tau_1), \ldots, \varepsilon(\tau_n)\rangle)$
- $\mathfrak{v}_{\text{va}}^M(\neg\phi) = \rho_\neg(\mathfrak{v}_{\text{va}}^M(\phi))$
- $\mathfrak{v}_{\text{va}}^M(\phi \wedge \psi) = \rho_\wedge(\mathfrak{v}_{\text{va}}^M(\phi), \mathfrak{v}_{\text{va}}^M(\psi))$
- $\mathfrak{v}_{\text{va}}^M(\phi \vee \psi) = \rho_\vee(\mathfrak{v}_{\text{va}}^M(\phi), \mathfrak{v}_{\text{va}}^M(\psi))$
- $\mathfrak{v}_{\text{va}}^M(\forall x\phi) = \begin{cases} \top & \text{if } \mathfrak{v}_{\text{va}_d^x}^M(\phi) = \top \text{ for all } d \in D \\ \bot & \text{otherwise} \end{cases}$
- $\mathfrak{v}_{\text{va}}^M(\exists x\phi) = \begin{cases} \top & \text{if } \mathfrak{v}_{\text{va}_d^x}^M(\phi) = \top \text{ for some } d \in D \\ \bot & \text{otherwise} \end{cases}$

From here we can define, e.g. D^+-validity in the usual way: If Γ and Δ are sets of sentences, then $\Gamma \vDash \Delta$ iff for every model M and variable assignment va, if $\mathfrak{v}_{\text{va}}^M(\phi) = 1$ for all $\phi \in \Gamma$, then $\mathfrak{v}_{\text{va}}^M(\psi) = 1$ for some $\psi \in \Delta$.

Before moving on to discussing rivalry, it's worth pausing to, again, see how to understand these things in terms of the model-theoretic account of logics given in Chapter 1. Many of the details are already visible, but since a few items require a bit of repackaging in order to fit that mould, it's worth taking our time to go through it explicitly.

W is a singleton set, e.g. . This corresponds to the intuition that each model of first-order logic is a "picture" of a way *the* world (and not, say, the *worlds*) could be.

$\mathcal{R}$ is not mentioned in any of the semantic clauses, so can be anything at all without impacting matters.

$\mathcal{D}$ is the function mapping to the domain of the model.

δ is exactly what above we called ε.

$\mathcal{V}$, .

$\mathcal{K}$, .

$\mathcal{C}$ is the collection consisting of the truth-functions ρ , ρ , ρ and the functions ρ and ρ , both of which are of the form $\mathcal{V}$ W $\mathcal{V}$ and are given by the piecewise definitions above.

$\mathfrak{v}(\Pi^n)$ is then defined exactly as it is above.

Gaps and Gluts, Again

The above formulation gets us first-order *classical* logic. If we want to allow gluts (for first-order LP) or gaps (for first-order K3) or both (for first-order FDE, the changes required are quite straightforward:

The function $\mathfrak{v}$, rather than having range , has whichever of , , b , , , n or , , b, n is appropriate.

For each 0-order connective c, the truth-function ρ_c is replaced with whichever of ρ_c^{LP}, ρ_c^{K3} or ρ_c^{FDE} is appropriate.[5]

For the quantifiers (first-order connectives), things are somewhat more complicated. It's best to explain it as follows: using language we introduced previously, we say that ϕ is *at least true* if $\mathfrak{v}(\phi)$ is either or b, and say that ϕ is *at least false* if $\mathfrak{v}(\phi)$ is either or b.

Then $x\phi$ is at least true if ϕ is at least true on each valuation va_d^x, and $x\phi$ is at least false if ϕ is at least false on some valuation va_d^x. Dually, we say $x\phi$ is at least true if ϕ is at least true on some valuation va_d^x, and $x\phi$ is at least false if ϕ is at least false on each valuation va_d^x. We then define $\mathfrak{v}(\ x\phi)$ (resp. $\mathfrak{v}(\ x\phi)$) to be iff $x\phi$ (resp. $x\phi$) is at least true but not at least false; iff $x\phi$ (resp. $x\phi$) is at least false but not at least true; b iff

[5] Recall that we did not define the truth-functions ρ_c^{FDE} explicitly, instead leaving this as an exercise for the reader.

$\forall x\varphi$ (resp. $\exists x\varphi$) is both at least true and at least false; and n iff $\forall x\varphi$ (resp. $\exists x\varphi$) is neither at least true nor at least false.

It is instructive to see what the above logics look like in terms of the extensions and antiextensions we introduced above. In the logics that lack gaps (e.g. classical logic and LP) E^+ and E^- obey the following constraint:

Exhaustion Constraint: For every n-ary predicate $^{3\ n}$, $E^+(^{3\ n}) \cup E^-(^{3\ n}) = D^n$.

In the logics that lack gluts (e.g. classical logic and K3) E^+ and E^- obey the following constraint:

Exclusion Constraint: For every n-ary predicate $^{3\ n}$, $E^+(^{3\ n}) \cap E^-(^{3\ n}) = \emptyset$.

So we have four first-order logics already at our fingertips and thus a bit of rivalry already. But this rivalry is, in some sense, parasitic on the rivalries we already discussed. That is, other than the fact that it's now dressed up in talk of extensions and antiextensions, nothing about the rivalry discussed here depends in any interesting way on our having moved to the setting of first-order logic. So we'll leave aside any further discussion of this particular sort of rivalry, and move on to other options.

Genuinely First-Order Rivalry

We'll now have a brief look at some strictly first-order options. To motivate the first one, we'll first prove that the following inference is D^+-valid on any of the four theories currently on the table:

$$W^1 p \pm \exists x W^1 x$$

Proof Suppose that for some M and va, $v^M_{va}(W^1 p) = \pm$. So $W^1 p$ is at least true on M and va. So $\exists x W^1 x$ is at least true. It follows that the inference is D^+-valid. □

But there is strong reason to think that this inference ought not be endorsed. Here is why: suppose that W^1 is the unary predicate ' x is a winged horse' and p is the individual constant 'Pegasus'. Since Pegasus in fact is a winged horse, it seems like $W^1 p$ is true. But then, since the above

3 Is Second-Order Logic Proper Logic?

In the previous chapter, we looked at various challenges to standard first-order logic, FOL for short. Many of the challenges to FOL have to do with theorems or inference rules of FOL that seem objectionable for some reason or other. An alternative to FOL would then try to avoid the inference or the theorem. Of course (as we also have seen in examples in Chapter 2), one might be discontent with FOL because of its expressive limitations and believe that there is more to *logic* than what is represented or modelled by FOL.

Let us begin with a simple example: let us assume that you started with FOL without adding '=' to its primitive vocabulary,[1] and you have convinced yourself that FOL as such does codify inferences and theorems that do count as logical. Now you consider adding '=' to that system (with the familiar semantic clauses and inference rules). What kind of considerations should matter for your decision? Of course the new language, $\mathrm{FOL}_=$, will lead to new theorems and will count more arguments as logically valid than mere FOL did. For example, the inference

1. The morning star is the evening star.
2. The morning star is a planet.
3. Hence the evening star is a planet.

can be shown to be valid, if (1) is formalized as '$a = b$', (2) as 'Pa', and the conclusion as 'Pb' (example is from Halbach, 2010). In FOL without identity, this inference could not have been shown to be valid, because (1) would have been formalized with a non-logical relation-letter 'R', supposedly formalizing the identity relation. But 'R' would, in different interpretations,

[1] $\mathcal{F}$ of the previous chapter had the identity sign already added as a logical constant. However, many versions of FOL that are taught in introductory courses do not initially contain identity.

have been mapped on other sets of ordered pairs than the identity relation. Hence there would be interpretations in which the inference would have a countermodel.[2]

So far, so good. The point of adding new constants would primarily be to increase expressive power in that sense. But then why not also add a new constant for 'is larger than'? This way, the obviously valid inference

1. The evening star is smaller than Uranus.
2. Uranus is smaller than Saturn.
3. Therefore the evening star is smaller than Saturn.

could also be shown to be valid in, say, $\text{FOL}_{=,\leq}$ (example again from Halbach, 2010), or how about adding operators $\ulcorner\Box\phi\urcorner$ for $\ulcorner$it is necessary that $\phi\urcorner$ and $\ulcorner\Diamond\phi\urcorner$ for $\ulcorner$it is possible that $\phi\urcorner$?

Some of the questions here concern the problem of how we should make a principled demarcation between logical constants and non-logical expressions. We will discuss this problem in detail in the next chapter. But instead of looking at the individual constants, one could also look at the logic that contains these constants as a whole and ask of that system whether this is still a logic.

Let's say we are convinced that FOL is a logic. Is $\text{FOL}_=$ a logic? $\text{FOL}_=$ allows us to express natural language sentences like 'The number of apples on this table is 3' in formulas that do not contain any "non-logical" terms other than those for 'apple', 'table' and 'is on', and, in particular, no terms for numbers.[3] Does that show that the natural language sentence really didn't refer to numbers, or does it show that $\text{FOL}_=$ has hidden mathematical content? In other words, does $\text{FOL}_=$ commit us to new entities over and above those we already recognized for FOL? What about $\text{FOL}_=$'s metalogic? Does the system have the same neat properties as FOL had, or does it have new ones that might disqualify it as a logic?

Such questions might matter for several reasons. You might just want to know what the limits of logic are. Where to draw the line between

[2] For example, a domain with two objects $\{1, 2\}$, $\mathfrak{I}(a) = 1$, $\mathfrak{I}(b) = 2$, $\mathfrak{I}(P) = \{1\}$ and $\mathfrak{I}(R) = \{\langle 1, 2\rangle\}$.

[3] If we take 'P' for 'is an apple' and 'Q' for 'is a table' and 'R' for 'is on', then the sentence can be formalized as $\exists x\exists y\exists z((Px \wedge Py \wedge Pz) \wedge (\neg x = y \wedge \neg x = z \wedge \neg y = z) \wedge \exists x_1((Qx_1 \wedge Rxx_1 \wedge Ryx_1 \wedge Rzx_1) \wedge \forall x_2((Px_2 \wedge Rx_2x_1) \rightarrow (x_2 = x \vee x_2 = y \vee x_2 = z))))$.

the logical and the non-logical. But you might also be interested in drawing that line for other purposes. As we shall see in Chapter 10, logicism is a view in philosophy of mathematics that holds that mathematics is nothing but logic (which is then supposed to explain some of the puzzling features of mathematics). But in order to show that mathematics is nothing but logic, one would first need to know how far logic extends.

Now, in the case of $\mathrm{FOL}_=$ we can put our minds at rest (at least that is the consensus). Going from FOL to $\mathrm{FOL}_=$ does not lead to drastic changes in the metalogical properties. The same is not true for other *extensions* of FOL. Modal logic has been heavily criticized (in particular by Quine (1953b, 1960)) for leading to new and problematic ontological commitments.[4]

Another such discussion has in recent years been concerned with the status of second-order logic, SOL for short. In this chapter we will look at the arguments for and against considering second-order logic to be proper logic. Before we can do so, we will give you a brief introduction to second-order logic(s).

Second-Order Logic(s)

Obtaining second-order logic from first-order logic is quite straightforward. In FOL we have basically two kinds of non-logical expressions: predicates and names.[5] FOL also has quantifiers and variables, and these variables syntactically take the same place as names. In other words, in FOL we are "quantifying into" name position. Thus, one of the two types of non-logical expression in the language, names, can be replaced by variables; the other, predicates, can't in FOL.

This changes in SOL. SOL also has variables for predicates. These predicate variables behave syntactically just like predicates, i.e. they come with different arities (places) and combine with terms (names or variables) in the same way that first-order predicates do. The language also

[4] For a discussion of Quine's arguments against modal logic, see Haack (1978), chapter 10.

[5] We are considering FOL without function symbols here.

has two new quantifiers, for universal and for existential quantification. The difference is that they now combine with predicate variables and quantify "into" predicate position. This results in formulas like the following:

(3.1) $\forall XXa$

(3.2) $\exists Y\forall x(Ya \rightarrow Fx)$

(3.3) $\forall X\exists yXy$

Predicate variables combine with terms to produce well-formed formulas. An n-place predicate variable must be followed by n-many terms in order to be well formed. Syntactically this is all as expected. We introduce a variable for the other type of non-logical expression, and that variable functions syntactically just as the type of expression it serves as a variable for.

If we like function letters in our first-order language (which allow us to construct complex singular terms, such as 'the father of ...'), we can also now add function *variables* to the language (again, if we like, with different arities), and quantifiers for these. The syntax of SOL doesn't require a primitive identity relation. Instead, with our new devices we can simply add identity by definition:

$$x = y =_{df} \forall X(Xx \leftrightarrow Xy)$$

The development of so-called "standard semantics" also looks like a straightforward and innocent extension of FOL. In FOL the interpretation function assigns objects from the domain to names and functions from the n-fold product of the domain to $\{\top, \bot\}$ to each n-ary predicate. In the last chapter, this was done via the functions δ and $\mathfrak{v}$. A variable assignment function, on the other hand, takes care of the individual variables, by basically doing what the interpretation function does for names: while the interpretation function assigns an object from the domain to every name, a variable assignment assigns to every name variable an object from the domain.

Let us simplify the semantics for FOL of the last chapter a bit, to see in what way we need to adjust the semantics in order to obtain SOL. Let us consider a single interpretation function $\mathcal{I}$ that combines the jobs of δ and $\mathfrak{v}$, assigning objects to names and functions of the appropriate sort to predicates. A model $\mathcal{M}$ is a tuple of the domain and the interpretation function, $\langle d, \mathcal{I}\rangle$.

Now in SOL, where we have predicate variables in addition to name variables, we need to take care of these. Since the variable assignment function is doing – in a manner of speaking – half the job of the interpretation function already, we can just have it do it all: in addition to assigning objects from the domain to name variables, in SOL we have the variable assignment function also assign functions $d^n \rightarrow \{\top, \bot\}$ to n-place predicate variables. In this case, the models of SOL are the same as those in FOL; they are ordered pairs of a domain and an interpretation function, $\langle d, \mathcal{I} \rangle$. What needs to be redefined is the definition of the variable assignment function (as just sketched), and the definition of the valuation function. Here are the two clauses we need to add to the latter (if we do not add function letters to the language):

- If Π is an n-place predicate variable and $\tau_1, \ldots, \tau_n$ are terms, and ε the result of glueing together $\mathcal{I}$ and va, then $\mathcal{M}, \text{va} \vDash_1 \Pi\tau_1 \ldots \tau_n$ iff $\langle \varepsilon(\tau_1), \varepsilon(\tau_2), \ldots, \varepsilon(\tau_n) \rangle \in \mathcal{I}(\Pi)$.
- If Π is a predicate variable and ϕ is a wff, then $\mathcal{M}, \text{va} \vDash_1 \forall\Pi\phi$ iff for every set U of n-tuples from d, $\mathcal{M}, \text{va}^{\Pi}_{U} \vDash_1 \phi$

(where va^{Π}_{U} is the variable assignment just like va except in assigning U to Π).

The definitions for *satisfiability*[6] and the definition for *semantic consequence*[7] remain just as they were for FOL. If we define the semantics for SOL like this, we obtain what is called the "standard semantics" for SOL. Since the interpretation function takes the semantic values for predicates from the set of all sets of n-tuples that can be generated from the domain (as it does in FOL), the predicate variables also range over that whole set.

In terms of the model-theoretic view of logics given in Chapter 1, almost everything is as it was in FOL with three exceptions:

- $\mathcal{C}$ now contains a truth-function (defined basically as above) for second-order quantification,
- δ is expanded as above, and (most interestingly)

[6] A formula ϕ is *satisfiable* iff there is a model $\mathcal{M}$ and a variable assignment va, such that $\mathcal{M}, \text{va} \vDash_1 \phi$.

[7] A formula ϕ is a *semantic consequence* of a set of formulas Γ iff every model that satisfies all formulas in Γ also satisfies ϕ.

- the function $\mathcal{D}$ now maps $\bullet$ to a pair consisting of the first-order domain d and the following set of functions:

$$\bigcup_{i=1}^{\infty}\{f : d^i \to \{\top, \bot\}\}$$

The thing to note, though, is the following: there's no obvious reason we *have to* include all of these functions into the second-order domain of quantification! After all, we are free to choose the first-order domain of quantification as we like; what's to keep us from similarly choosing the second-order domain as we like as well?

If we do this – if we allow ourselves to specify separately what the domain of the second-order variables is, rather than simply taking it to contain all of the available functions – the result is what are known as *Henkin models*.

More explicitly, a Henkin model is a triple $\mathcal{M}^{\mathrm{H}} = \langle d, d^*, \mathcal{I}\rangle$ where d is the first-order domain and $\mathcal{I}$ is an interpretation function as above, but d^* is a subset of the set $\bigcup_{i=1}^{\infty}\{f : d^i \to \{\top, \bot\}\}$. The variable assignment function now assigns members of d to each individual variable, but appopriate members of d^* to each predicate variable. The definition of the valuation function can basically remain as it was, replacing $\mathcal{M}^{\mathrm{H}}$ for $\mathcal{M}$.

Henkin models can be equivalent to standard models. That's the case if d^* contains all the semantic values that a standard model would have provided as semantic values for the predicate variables (rather than a proper subset of them). A Henkin semantics that would be restricted to such models would be equivalent to a standard semantics. This gives us the following results about the relationship between these two semantic theories (see Shapiro, 1991):

- If ϕ is valid according to Henkin semantics, then ϕ is valid according to the standard semantics.
- If ϕ is a semantic consequence of Γ according to Henkin semantics, then ϕ is a semantic consequence of Γ according to the standard semantics.
- If ϕ is satisfiable according to standard semantics, then ϕ is satisfiable according to Henkin semantics.

None of the converses hold.

Now, we can also construct a deductive system for SOL. As usual, this can be provided in different styles, be it axiomatic (Shapiro, 1991), natural deduction (Prawitz, 1965), or tableaux (Jeffrey, 1967).

The system that we introduced in Chapter 2 would just need to be adjusted by adding rules for the second-order quantifiers. In the rules below, 'Φ' and 'Ξ' are open sentences, 'F^n' is an n-place predicate, and 'X^n' is an n-place predicate variable:

Universal Quantifier

$$\frac{\Gamma, \Phi(\Xi^n/X^n) \Rightarrow \Delta}{\Gamma, \forall X^n \Phi \Rightarrow \Delta}\ \forall^2\mathrm{L} \qquad \frac{\Gamma \Rightarrow \Phi(F^n/X^n), \Delta}{\Gamma \Rightarrow \forall X^n \Phi, \Delta}\ \forall^2\mathrm{R}$$

Existential Quantifier

$$\frac{\Gamma, \Phi(F^n/X^n) \Rightarrow \Delta}{\Gamma, \exists X^n \Phi \Rightarrow \Delta}\ \exists^2\mathrm{L} \qquad \frac{\Gamma \Rightarrow \Phi(\Xi^n/X^n), \Delta}{\Gamma \Rightarrow \exists X^n \Phi, \Delta}\ \exists^2\mathrm{R}$$

We need to observe the *eigenvariable restrictions* as they were introduced in Chapter 1 (p. 32). SOL with the standard semantics discussed above and the deductive system just introduced has some fascinating metalogical properties. Since they have been of relevance in the discussion of SOL's status, we will explain a short selection of these.

Let us begin with the perhaps less exciting side. If we consider SOL with the Henkin semantics, $\mathrm{SOL}^{\mathrm{H}}$, the following metalogical results for FOL all carry over to $\mathrm{SOL}^{\mathrm{H}}$ (here presented in their FOL formulation (Shapiro, 1991, 79–80)):

Soundness Let Γ be a set of formulas and ϕ a formula of FOL. If $\Gamma \vdash \phi$ then $\Gamma \vDash \phi$ (every formula derivable from Γ is a semantic consequence of Γ). *A forteriori*, if $\vdash \phi$ then $\vDash \phi$.

Completeness Let Γ be a set of formulas and ϕ a formula of FOL. If $\Gamma \vDash \phi$ then $\Gamma \vdash \phi$ (every semantic consequence of Γ is derivable from Γ). *A forteriori*, if $\vDash \phi$ then $\vdash \phi$.

Compactness Let Γ be a set of formulas of FOL. If every finite subset of Γ is satisfiable, then Γ is satisfiable.

Downward Löwenheim–Skolem theorem If $\mathcal{M}$ is a model of a set Γ of FOL formulas, then $\mathcal{M}$ has a submodel $\mathcal{M}'$ whose domain is at most countably infinite, such that for each assignment s on $\mathcal{M}'$ and each formula ϕ in Γ: $\mathcal{M}, s \vDash \phi$ if, and only if, $\mathcal{M}', s \vDash \phi$.

Upward Löwenheim–Skolem theorem Let Γ be a set of FOL formulas. If, for each natural number n, there is a model of Γ whose domain has at least

n members, then for any *infinite* cardinal κ, there is a model of Γ whose domain has cardinality at least κ.

These results might all be familiar to you from studying FOL. *Soundness* is the minimal requirement we expect from a logic: if we can prove a formula from a set of premises, that formula should also be a semantic consequence of those premises. *Completeness* is the more ambitious requirement that every consequence can indeed be reached by proof. Both concern the relation between a specific model theory (in case of FOL, the standard semantics; in case of SOL^{H}, the Henkin semantics) and a specified proof system. *Compactness*, *Downward* and *Upward Löwenheim–Skolem* are perhaps less familiar. However, one might consider these results to represent the expressive limitations of FOL (and thus of SOL^{H}).

Compactness means that "you can't say anything in [FOL] whose logical significance would emerge only in connection with infinitely many other sentences" (Sider, 2010, 105). Remember that we mentioned above that $\text{FOL}_{=}$ allows us to express numerical claims, such as 'there are three apples on the table'. Even though FOL can express these numerical statements, it can't express claims like 'there are finitely many apples'.

The other two results show similar weaknesses of FOL. Upward Löwenheim–Skolem entails that every first-order theory with a countably infinite model, e.g. Peano Arithmetic, has an *uncountable* model, too. By the Downward Löwenheim–Skolem theorem, real analysis which has as the intended uncountable domain the real numbers has a countable model (for details, see Shapiro, 1991). In general, first-order theories with countable models also have uncountable models, and first-order theories with uncountable models also have countable models.

Interestingly, these expressive limitations vanish if we move to SOL with standard semantics. Compactness, Downward and Upward Löwenheim–Skolem all fail to hold for SOL. This allows SOL to formalize second-order theories that have *categorical* axiomatic systems for infinite structures. An axiomatic system is *categorical* iff all of its models are *isomorphic*.[8]

[8] There is a structure-preserving one-to-one mapping between the objects of these models.

On a less positive note, completeness also fails for SOL with standard semantics. Thus, not all semantic validities of SOL are provable.[9] This can actually be shown on the basis of the categoricity result (see Shapiro, 1991): Let *G* be the Gödel sentence[10] of the (finite) second-order axiom system of arithmetic. Dedekind has shown that if we replace the induction scheme by its second-order sentence, we get a categorical theory. Gödel has shown that *G* is true in the standard model, albeit not provable in the deductive system. Let us now consider the conjunction of the axioms of that system, *AR*. The sentence '$AR \rightarrow G$' can't be proven in the deductive system of SOL (for the reasons just mentioned). However, this conditional is valid in the standard semantics of SOL, as follows from the categoricity result for arithmetic.

Note that the problem of incompleteness is not just that there are some axioms missing, or some weird theorems not provable. The problem is rather that the consequence relation of the standard semantics is non-recursive and thus – no matter what you do – you can't have a (recursive) proof system.

The incompleteness of SOL has a further, somewhat ironic, dimension. We can formalize sentences in SOL, let's call them 'CH' and 'NCH', such that CH is true iff the continuum hypothesis[11] is true, and NCH is true iff the continuum hypothesis is false. Now, either NCH or CH is a validity of the standard semantics of SOL, and hence a logical truth of SOL (Shapiro, 1991, 105). The ironic part is that, although SOL *entails* the solution to an

9 We said above that completeness is a claim about a semantics and proof system. Saying that SOL is incomplete means that there is *no* proof system that would allow us to prove all validities of SOL.

10 Gödel's *First Incompleteness Theorem* says that for a formalized system *F* which contains Robinson arithmetic a sentence *G* of the language of *F* can be mechanically constructed from *F* such that if *F* is consistent (in a specific sense) then neither *G* nor its negation is derivable in *F*. Such an "undecidable" (that is, neither provable nor refutable in *F*) statement is then called "the Gödel sentence" of *F*. For details see Raatikainen (2018).

11 The *continuum hypothesis* says that there is no set whose cardinality is strictly between that of the integers and the real numbers. The hypothesis was first formulated by Georg Cantor in 1878. The hypothesis is independent of ZFC set theory (Zermelo–Fraenkel set theory with the axiom of choice, the considered "standard" version of set theory), i.e. the hypothesis or its negation can be added to ZFC and the resulting theory is consistent if ZFC is consistent.

open problem of mathematics (and there are more of this kind), SOL keeps these solutions to itself: because of SOL's incompleteness we can't prove CH or NCH.

Ontological Commitments

SOL seems to be a straightforward extension of FOL. FOL distinguishes two kinds (or, if we include function letters, three kinds) of non-logical expression and introduces quantifiers and variables for one of these. However, there are also natural language arguments and inferences that concern the other non-logical expressions, hence SOL introduces variables and quantifiers for these. The semantic values of the expressions over which the quantifiers now range are just those that were already assumed to be the semantic values of predicate letters (or function letters) in FOL (at least in standard SOL). Why not then consider SOL to be *logic*, if FOL clearly qualifies as such?

The most famous argument against considering SOL to be logic stems from Quine (1970). In this book, Quine argues that SOL is not logic, but in fact "set theory in sheep's clothing". Quine's argument is a bit hard to reconstruct. It's placed in the section *The Scope of Logic*, in which he also considers other candidates for extensions of logic. In fact, as in our chapter, the first candidate he considers is "identity theory", i.e. extending FOL to $\mathrm{FOL}_{=}$. Quine opts in favour of this extension, because it (a) retains completeness for FOL, (b) is universal (identity theory doesn't discriminate between different objects in the way a theory of numbers would), and (c) once the machinery for FOL is in place anyway, identity theory can be added without requiring any extra machinery.

Set theory, however, gets thumbs down. In the section that is supposed to have Quine's argument against set theory, he considers the question why some philosophers came to believe that it was logic. He speculates that a reason or at least a sign of that confusion has to do with SOL. In the section *Set Theory in Sheep's Clothing* he argues against three "confusions". The first is that second-order quantification is quantification over predicates, the second is to confuse the intension of predicates, i.e. attributes, with their extensions, i.e. sets. The third, finally, is supposed to be the confusion that second-order quantification would have to treat predicates to be names of their extensions, but predicates just aren't names. Here is

Quine's argument why variables eligible for quantification do not belong in predicate position:

> Consider first some ordinary quantifications: '$(\exists x)$(x walks)', '(x)(x walks)', '$(\exists x)$(x is prime)'. The open sentence after the quantifier shows 'x' in a position where a name could stand; a name of a walker, for instance, or of a prime number. The quantifications do not mean that names walk or are prime; what are said to walk or to be prime are things that could be named *by* names in those positions. To put the predicate letter 'F' in a quantifier, then, is to treat predicate positions suddenly as name positions, and hence to treat predicates as names of entities of some sort. The quantifier '$(\exists F)$' or '(F)' says not that some or all predicates are thus and so, but that some or all entities of the sort named by predicates are thus and so. (Quine, 1970, 7)

Quine seems to assume that since ordinary variables occur in name position, all variables must always occur in positions in which a name could occur. But that just seems to beg the question against SOL. As Boolos (1975) argues, we wouldn't have believed the following argument:

> Consider some extraordinary quantifications: '$(\exists F)$(Aristotle F)', '(F)(Aristotle F)', '$(\exists F)$(17 F)'. The open sentence after the quantifier shows 'F' in a position where a predicate could stand; a predicate with an extension in which Aristotle, for instance, or 17 might be. The quantifications do not mean that Aristotle or 17 are in predicates; what Aristotle or 17 are said to be in are things that could be had *by* predicates in those positions. To put the variable 'x' in a quantifier, then is to treat name positions suddenly as predicate positions, and hence to treat names as predicates with extensions of some sort. The quantifier '$(\exists x)$' or '(x)' says not that some or all names are thus and so, but that some or all entities of the sort had by names are thus and so. (Boolos, 1975, 510)

According to Boolos we should reject the last two claims of that argument as false and not following from the preceding ones, and for the same reasons reject the final claims of Quine's original argument (Boolos, 1975, 510).

Jason Turner (2015) reconstructs what he calls "Textbook Quineanism", the argument that most philosophers seem to have in mind when they believe that Quine has shown that SOL is committed to sets. The argument runs like this:

1. You are ontologically committed to something of a particular kind if and only if the bound variables in your system have to range over things of that kind to be true.
2. The bound variables of second-order logic have to range over something predicate-like.
3. Therefore, theorems of second-order logic ontologically commit you to something predicate-like.
4. So second-order logic is not logic. (Turner, 2015, 469)

As Turner points out, this argument doesn't seem to be very Quinean. Quine does not think that you can just read off the ontological commitments of a theory (as premise 1 states). Instead, the ontological commitments of a theory only appear *after* the theory is properly reconstructed. *Inter alia* this requires that the theory is translated into a first-order theory. Only after this is done can the ontological commitments be read off.

Perhaps that is indeed what Quine has in mind, when he argues that the higher-order claim $\ulcorner \exists X \forall x (Xx \leftrightarrow Fx) \urcorner$ follows from $\ulcorner \forall x (Fx \leftrightarrow Fx) \urcorner$, but only disguises the statement $\ulcorner \exists y \forall x (x \in y \leftrightarrow Fx) \urcorner$ which clearly is committed to a set (Quine, 1970, 68).

But there are at least two objections to the suggestion to translate second-order claims this way into first-order logic. (i) as Boolos argued, this translation does not preserve validity or implication. $\ulcorner \exists X \forall x Xx \urcorner$ is valid, while $\ulcorner \exists x \forall y (y \in x) \urcorner$ isn't. $\ulcorner \forall X (Xa \rightarrow Xb) \urcorner$ implies $\ulcorner a = b \urcorner$, but $\ulcorner \forall x (a \in x \rightarrow b \in x) \urcorner$ doesn't (Boolos, 1975, 512). This should speak against the adequacy of the suggested translation.

Also, one might question why Quine believes that only a first-order translation of a theory can properly capture the logical commitments of a theory. If Quine's answer to that question is that first-order logic is the only proper logic, then he is simply begging the question against SOL again. After all, SOL's putative commitment to sets is supposed to be the reason for not considering SOL to be proper logic (and thus the right framework to represent its ontological commitments).

But let's leave Quine's original argument and let's turn instead to Turner's reconstruction of "Textbook Quineanism". What to say about this argument? Premise 2 could be motivated by arguing that the use of a quantifier always ontologically commits. That's just what quantifiers mean. Now, since in SOL the variables that are quantified over are the kinds of things that can be used predicatively, they have to be predicate-like entities, like sets, or perhaps Fregean concepts (Turner, 2015, 470).

But this argument can be resisted. Here is a counterargument by A. N. Prior:

> Quine would argue, I think, that the quantified forms $\ulcorner\forall xFx\urcorner$ and $\ulcorner\exists xFx\urcorner$ do not commit us to the existence of any other sorts of entities than do the corresponding singular forms $\ulcorner Fa\urcorner$, $\ulcorner Fb\urcorner$, etc., which follow from the former and entail the latter. Why, then, should he suppose that the quantified forms $\ulcorner\exists XXa\urcorner$, $\ulcorner\exists X\exists xXx\urcorner$, etc., commit us to the existence of sorts of entities to which we are not committed by the forms $\ulcorner Fa\urcorner$, $\ulcorner Ga\urcorner$, $\ulcorner\exists xFx\urcorner$ from which they follow? [...] The alleged emergence of these new ontological commitments has an almost magical air about it. (Prior, 1971, 43)

Prior argues here that quantification into name position carries ontological commitment, because names carry such ontological commitment. In order for 'Socrates is wise' to be true, there must be a referent for 'Socrates'. Thus, this statement commits you to Socrates. If we believe that it doesn't also commit us to wisdom, but only to Socrates being a certain way, why should all of a sudden quantification into predicate position create such ontological commitment? As Prior says, this seems to be magical. Or, as Turner puts it, when we assert sentences like '$\exists X(X(Socrates))$' we are not saying that there is some predicate-like entity that Socrates participates in, any more than when we assert 'Socrates is wise'. In the latter we say how Socrates is particularly, in the former how he is in more generality (Turner, 2015, 471).

But perhaps the problematic commitment to sets isn't part of the *object language*, but rather comes with its model-theoretic semantics. After all, we seem to make use of sets when setting up the standard semantics of second-order logic. Indeed, the model theory is committed to sets, but is it committed to sets in a problematic way? Boolos (1975, 48) already observed that the commitment of second-order logic is not comparable to full set theory. Marcus Rossberg summarizes this observation as follows:

> Generally, as the second-order variables are interpreted as ranging over the subset of the domain, we will end up with a "second-order ontology" of the size of the powerset of the first-order domain. If there are n objects in the domain, the second-order variables range over 2^n sets. Only one application of the powerset axiom is needed, though, which is not very much

> compared to the vastness of the set-theoretical hierarchy. So, speaking in set-theoretical terms, the commitments of second-order logic are still fairly modest [. . .]. (Rossberg, 2006, 49)

But there are more possible moves. One could be to separate *model theory* and *semantic theory* and consider the model theory to provide a model of the consequence relation, but not to provide the proper truth-conditions of the language (Turner, 2015). Another would be to note that also first-order model theory talks of sets when explaining the truth of sentences like '*Fa*', but we don't think that the model-theoretic semantics is ontologically committing, if we don't think that the original sentence was; the same should then hold for second-order model theory.

Finally, one could attempt to provide the semantics in other terms. Perhaps the semantics for second-order logic should not be provided in a first-order metalanguage. Agustin Rayo and Tim Williamson provide instead a second-order metalanguage (Rayo and Williamson, 2003).

Plural Quantification and Intelligibility

An alternative argument against SOL, which is also perhaps the best contester for a real Quinean argument against SOL, is what Jason Turner called the *Intelligibility Argument.*

1. A formal system is meaningless unless it is provided an interpretation – a specification of what its expressions mean.
2. We provide an interpretation by specifying the meaning of each expression using terms already understood.
3. The only plausible meaning for '$\exists X \ldots$' is 'There is a predicate-like thing X such that . . . '.
4. So, if second-order logic is meaningful, the second-order quantifiers say that there are predicate-like things.
5. If the second-order quantifiers say that there are predicate-like things, then second-order logic is not logic.
6. Therefore, either second-order logic is meaningless, or it is not logic. (Turner, 2015, 473–474)

Perhaps this could serve as the missing part in Quine's argument that we discussed above. Maybe Quine isn't just presupposing the idea that variables *qua* variables have to occur in name position and that therefore

SOL is quantification over predicate-like things, but rather (perhaps implicitly) arguing that that's the only way to make sense of quantification.

This argument might be resisted in one of two ways: either one attacks the idea that we always have to provide an interpretation in terms already understood, or one attacks the idea that there is only one way to understand second-order quantifiers.

The first strategy would challenge premise 2. As Turner (2015) argues, the demand that we always have to interpret formal systems in familiar terms is pretty restrictive and not in line with usual scientific practice. Often we introduce *theoretical terms* by implicit definition, i.e. we put down axioms in which the new expressions feature together with expressions already understood. If there is something in the world that satisfies these axioms, then that will be the extension of the theoretical terms so introduced (if there's nothing that makes the theory true, then the new terms just fail to have an extension). But this strategy might fail in this case. First of all, because of the incompleteness of SOL, we won't be able to fix an interpretation for SOL by writing down a set of axioms. Moreover, the agnosticism built into this strategy invites a skeptical response: what if the unique satisfier of the theory is *sets* after all?

This might make it more promising to look for a strategy that could challenge premise 3. An early attempt of that kind is Susan Haack's substitutional interpretation of SOL (Haack, 1978). According to such a substitutional interpretation, a quantified statement is true iff there are expressions of the suitable syntactic category in the language such that the quantified expression with the variables substituted with these expressions come out true. For the case of second-order quantification, a formula $\ulcorner \exists X\phi(X) \urcorner$ is true iff for some predicate Π, $\ulcorner \phi(\Pi) \urcorner$ is true. The latter can be the case without committing us to predicate-like entities. 'Socrates is wise' does not commit us to a predicate-like entity. The problem with such an account is that it doesn't actually say what the *meaning* of $\ulcorner \exists X\phi(X) \urcorner$ is, but merely states a *truth-condition*, and these, arguably, aren't the same (Turner, 2015). If, on the other hand, we want to identify meaning and truth-condition, we end up saying that $\ulcorner \exists X\phi(X) \urcorner$ *means* 'there is a predicate such that ...', which, of course, commits to predicate-like things (namely, predicates).

Another strategy of this kind is Boolos' observation that monadic SOL can be translated into first-order *plural logic* (Boolos, 1984). Also independently of this observation plural quantification has received a lot of recent attention. Let's quickly introduce the main idea behind it (using the system from Linnebo (2014)).

The syntax of the language L_{PFO} is an extension of FOL. In addition to our singular variables (x, y, etc.), we will also have "plural variables" (xx, yy, etc.) and in addition to singular name constants (a, b, etc.) we will now also have plural constants (aa, bb, etc.). We have two predicate constants, the familiar '$=$' for *identity* and the new '$\prec$' for 'is one of' that can stand between a singular and a plural term to form a well-formed formula. Since we have plural variables, we will have two plural quantifiers $\forall vv$ ('for any things') and $\exists vv$ ('for some things'), where vv is a plural variable. In this language, we can express the famous Geach–Kaplan sentence,

(GK) Some critics admire only one another.

as

(GK′) $\exists xx[\forall u(u \prec xx \rightarrow Cu) \wedge \forall u \forall v(u \prec xx \wedge Auv \rightarrow v \prec xx \wedge u \neq v)]$

The idea is that L_{PFO} captures a form of quantification that we find in ordinary language, such as the Geach–Kaplan sentence above, that can't be properly paraphrased in FOL. Linnebo (2014) describes the details of how to build a theory *PFO* of first-order quantification based on the language L_{PFO}.

Boolos argued that plural quantification is a familiar feature of ordinary language. Moreover monadic second-order logic, MSOL and PFO are *equi-interpretatble*; their theorems can be mapped onto each other. Does that show that there is an alternative, ontologically unproblematic interpretation of SOL?

This will depend on a variety of factors. First, it will of course depend on the status of PFO itself: is PFO itself proper logic? Are its principles (e.g. the introduction and elimination rules for the new quantifiers) logically valid? Is PFO ontologically innocent?

On the last questions, Boolos was quite confident that plural quantifications in ordinary language do not carry problematic ontological commitments. (GK) just talks about critics that admire other critics, but not about sets or other "plural entities". If so, then a plural predicate such as '*Fxx*' can be jointly satisfied (or unsatisfied) by several individuals at once. This view is called "pluralism" (and contrasts with "singularism", the view that such predicate is satisfied or unsatisfied by a single individual) (Rayo, 2007).

Some philosophers do not share Boolos' intuitive reading of (GK) (cf. Resnik, 1988) and argue that already the natural language sentence quantifies over *subsets* of the critics.

Perhaps most importantly, it is not clear that establishing that there is a mutual interpretability of MSOL and PFO establishes that MSOL (let alone SOL) is unproblematic *even if we assume that PFO is unproblematic*.

> [Equi-interpretability] does not show [by itself] anything about these two pairs of theories' being equivalent in any of the more demanding senses that philosophers often care about (such as having the same epistemic status, ontological commitments, or degree of analyticity). (For instance PFO is equi-interpretable with atomic extensional mereology, which philosophers tend to find much more problematic than PFO.) In order to show that the two pairs of theories are equivalent in some philosophically important respect *F*, we would need to show that the above translations preserve *F*-ness. (Linnebo, 2014)

Plural quantifiers can't be simply interpretations of second-order quantifiers either. There is no analogue of dyadic second-order quantification (which is why Boolos' project is confined to MSOL), and the plural variables don't take predicate position (which was the obstacle to interpreting SOL) (cf. Rayo, 2007).

But there is a further project, carried out by Rayo and Yablo (2001), which is similar in spirit to Boolos' project, but makes use of another feature of ordinary language. They depart from the observation that *pro-adverbs* such as 'likewise' or 'somehow' allow for non-committal readings of second-order sentences with polyadic predicates. For example (see Turner, 2015, 475),

$\exists X(X(Scooby, Shaggy))$

should not be understood as saying 'There is a way Scooby and Shaggy are related', because that would be to quantify over ways, but instead understood as saying

> Scooby and Shaggy are somehow related.

This would deal with the issues we raised for Boolos' project: the approach can deal with polyadic SOL and the plural variables don't take name positions. However, this approach has to fight with technical difficulties of its own, and with whether the ordinary language translations really deliver the results that full SOL requires (cf. Rossberg, 2015; Turner, 2015). To consider an example from Turner (2015, 476), the SOL formula

$$\exists X \forall x \forall y \forall z (X(x, y, z) \leftrightarrow x = \mathit{Scooby} \land y = \mathit{Shaggy} \land z = \mathit{Velma})$$

will be true on all full second-order models, because there is a set containing just one triple with these three objects of the domain that are assigned as referents. However, on the reading suggested by Rayo and Yablo (2001) this becomes

> Things somehow relate such that any three things that are so related if and only if the first is Scooby, the second Shaggy, and the third Velma.

But it's not clear that this last sentence really says that there is a relation that is had only by Scooby, Shaggy and Velma, in that order (and not also had by three other things).

Standard Semantics and Mathematical Content

The arguments we have looked at so far applied to SOL in general and didn't just find fault with *full* SOL, i.e. SOL with *standard semantics*. In this and the next section we will focus on the latter.

When introducing standard semantics for SOL we noticed two features which might be considered an obstacle to SOL being proper logic. SOL is incomplete, and SOL contains validities the truth-value of which varies together with the truth-value of certain mathematical hypotheses. In particular, there are sentences of SOL, CH and NCH, that are truths of SOL iff the continuum hypothesis is true, false, respectively.

Now that is mathematical content *right there*. Who cares about whether sets are somehow disguised behind quantification into predicate position? SOL entails mathematical, set-theoretical claims among its theorems!

As Turner (2015) explains, there are at least two considerations against a consequence relation that treats CH or NCH as validities. First of all, such a consequence relation would violate the *topic neutrality* of logic. Whether a sentence follows from a set of premises and whether a sentence is a logical validity should not depend on the truths of any specific subject matter. It should hold in complete generality. But the validity of CH or NCH depends on the truths of set theory in the way specified above: we can prove that CH is a validity of SOL iff the continuum hypothesis is true.

Note, however, that there are validities of first-order logic that "depend" in a similar way on set-theoretical truths. First of all, there is the first-order inference

$$\frac{\begin{array}{c}\forall x(Fx \to Gx)\\ \forall x(Gx \to Hx)\end{array}}{\forall x(Fx \to Hx)}$$

which is valid iff the subset relation is transitive (cf. Turner, 2015). Secondly, there is the first-order statement

$$\forall x[x \neq S(x) \wedge 0 \neq S(x) \wedge \forall y \neq x(S(y) \neq S(x))]$$

which is unsatisfiable if there are no infinite sets (because then there are no infinite domain models). Thus, the negation of the above statement is a first-order validity unless there are sets with infinitely many members (cf. Rayo and Yablo, 2001).

So, unless one is prepared to also count first-order logic to be "set theory in sheep's clothing", SOL doesn't seem to be particularly objectionable.

But there is a second way to build an argument against SOL's logicality on CH and NCH. One might consider it part of logic's *normativity* that we have certain rational obligations towards logical validities. In particular, one might hold that it is an error of reasoning to reject genuine validities. However, it doesn't seem to be an error of reasoning to reject either CH or NCH; they seem to be "radically epistemically unsettled" (Turner, 2015, 480).

But this "deep epistemic openness" of CH and NCH doesn't primarily have to do with their (indirect relation to) set-theoretic content, but rather with the incompleteness of full SOL. This is a metalogical result that makes SOL radically different from FOL. Doesn't that suffice to establish that SOL isn't logic?

Standard Semantics and Incompleteness

Is the mere fact that SOL with standard semantics is incomplete a reason against considering SOL to be proper logic? Let's remind ourselves again what a completeness proof establishes. We have two mathematical theories, if you like: one we can call the deductive system, and the other we call the semantics. As Dummett (1978) explains, there is a technical interest in knowing that the semantics is sound and complete with respect to the deductive system. Knowing of soundness and completeness allows us to derive conclusions about the deductive system from results in the semantics and vice versa. If we can show that there is a model that satisfies a set of sentences Γ and, at the same time, falsifies the sentence ϕ, we know that ϕ can not be derived from Γ. Likewise, we can sometimes establish logical consequence "faster" by reasoning through the semantics than by reasoning through the deductive system (cf. Rossberg, 2004, 311). As Dummett notes, for this technical purpose it isn't actually necessary that the second of these two systems, which we called the semantics, really is a semantics for the language. Any "algebraic devise involving functions defined over a two-element set" can do the job, if its classification of formulas corresponds to the deductive system in the appropriate way (Dummett, 1978, 294).

The situation changes a little if we have reason to put special trust in one of the two systems. Soundness and completeness proofs can then lead us to trust the other system as well (cf. Rossberg, 2004). Also, *if* we have already convinced ourselves of the justifiedness of a system, the soundness and completeness proofs can deliver an internal explanation of why the system works (cf. Dummett, 1978). Is there a guarantee that such an explanation is always available and a principled argument that *logic* has to deliver it? It's not clear that this is the case. The soundness result for SOL might provide us with an explanation for why our second-order inferences are trustworthy, but perhaps there is no other assurance of our semantic reasoning than knowing that it is a (relatively) straightforward extension of our first-order semantic reasoning.

Let's focus on *completeness* in particular. Why should the completeness of a system matter for its value as a logic? There are considerations, which we will discuss in more detail in Chapter 8, that could perhaps offer an answer to this question. Logic is supposed to have a certain normative authority on our reasoning. In other words, the fact that ϕ follows from a set of premises Γ will have some normative consequence on how we may rationally reason, if we believe all sentences in Γ, say. But even if this normative consequence only holds in an idealized way (for example, under conditions of "ideal reflection" etc.) we might nevertheless require that knowledge of logical consequence or of logical truth must be humanly possible in order to have such normative force. If it is not knowable that ϕ follows from Γ, it is unclear why this fact should have any normative consequence, however indirect, on what we ought to believe. This consideration now connects to the idea we discussed at the end of the last subsection: if CH and NCH are "radically epistemically unsettled", they can't play the normative role we expect of logical truths.

And SOL with standard semantics just is "anything but accessible to the human mind" (cf. Raatikainen (Forthcoming) who shows that SOL also outstrips less demanding requirements than completeness, or the recursive enumerability of its logical truths). If we think that it is a central function of logic to play such a normative role (cf. Field, 2009a), then full SOL is not a logic. We will postpone the discussion of whether logic has any such normative role to Chapter 8.

There is a further consideration for why completeness proofs in particular might matter. Our initial puzzle arises because we see completeness proofs as merely establishing a result between two formal notions: model-theoretic validity and proof-theoretic derivability. As we said above already, this overlooks that we have reason to put some trust in these systems. As Georg Kreisel (1972) argues, leaving this out is an oversight that might be due to a positivistic prejudice to discard any informal notions.

Kreisel suggests considering three notions: that of the intuitive validity of a formula, that of validity in all formal set-theoretic structures, and that of formal derivability by means of some fixed set of formal rules. We can introduce three corresponding predicates. $Val\alpha$ for 'α is intuitively valid', $V\alpha$ for 'α is valid in all set-theoretic structures', and $D\alpha$ for 'α is formally derivable by means of an accepted set of rules'.

Kreisel argues that there are two accepted properties of *Val*. On the one hand, we accept – in mathematical practice – the validity of inference rules. For example, Frege's system was accepted long before there even was such a thing as model-theoretic semantics. Thus, we accept (where α^i denotes formulas of order i):

$$\forall i \forall \alpha (D\alpha^i \rightarrow Val\alpha^i)$$

Moreover, we also accept in mathematical practice that

$$\forall i \forall \alpha (Val\alpha^i \rightarrow V\alpha^i)$$

because we take logic to apply to mathematical structures, and hence to set-theoretic structures. So we have reassurance that the rules don't lead us astray *and* we have trust in the applicability of logic to mathematics. But then, Gödel's completeness result for first-order predicate logic, i.e.

$$\forall\alpha^1(V\alpha \rightarrow D\alpha)$$

allows us to infer two further facts, namely

$$\forall\alpha^1(V\alpha \leftrightarrow D\alpha^1)$$

and

$$\forall\alpha^1(V\alpha \leftrightarrow Val\alpha)$$

Thus, the completeness result allows us to *prove* the equivalence of our intuitive notion with our semantic and proof-theoretic notions of validity. Of course, Kreisel's proof rests on facts about *Val* that are not themselves formal results, but rather presuppositions of mathematical practice. His argument does show why completeness proofs are important when we have them (because they allow us to tie our formal and informal notions together). However, as Rossberg (2006) argues, this does not show that if we lack a completeness result (as in the case of second-order logic) the systems in question aren't logics.

Before we close the chapter we should add a comment on SOL's deductive strength. As we have seen, most of the features that make SOL attractive in the first place (for example, the categoricity results) depend on its strong model theory. However, the interest that *neologicism* has in SOL (as discussed in Chapter 10) connects to the strength of its axiomatic deductive system "only". As we explain in Chapter 10, SOL is a central part of the neologicist's project to derive certain mathematical theories from *Hume's Principle*. Hume's Principle taken by itself is relatively weak (in terms of interpretability of mathematical theories) however; in combination with the standard inferential apparatus of SOL it suddenly becomes very strong. Raatikainen takes this as an indication that the axiomatic deductive system, taken by itself, is not mathematically innocent either, and thus shouldn't qualify as logic proper.

The status of SOL seems still open, if you consider the question of logical status to have an objective answer. We looked mainly at Quinean reasons against considering SOL to be logic proper. However, one can also take to heart another lesson from Quine's writings and hold with Stewart Shapiro that the demarcation between logic and the rest of science (and mathematics in particular) is not strict. What we want to regard as *logic* is a matter of *pragmatic* choice, and thus a matter of first settling what we want a logic to do for us.

One of the tasks of logic is to codify the norms of *deductive* reasoning, a form of reasoning that is especially prevalent in mathematics. Thus, a logic should be able to adequately represent mathematical reasoning. As we have seen above, the adequate representation of mathematical reasoning and the content of mathematical theories requires *categoricity*. Therefore, an adequate "logic" should be of the strength of SOL with standard semantics.

Likewise, symmetry considerations might speak in favour of SOL. We can express in $FOL_=$ that there are n objects, but not that there are finitely many objects. Why shouldn't logic be able to express all cardinality quantifiers?

Moreover, above we have seen that Quine's arguments against generalization into anything other than name position are unconvincing. But then it just seems to be a violation of logic's *universality* that we restrict the syntactic places into which we can generalize in this manner.

Exercises

1. Why does *Compactness* entail that $FOL_=$ can't express 'There are finitely many *F*'? Can you prove this?
2. What is the closest translation to $FOL_=$ of the Geach–Kaplan sentence (GK) that you can come up with? How would the sentence look in SOL?
3. We argued that the fact that either CH or NCH is a semantic consequence of SOL could speak against SOL's logicality in virtue of logic's supposed *normativity*. Rejecting CH or NCH does not seem to be an error of reasoning, while rejecting a logical validity does. But is that so? What about non-classical logicians that reject validities of classical logic? Are they committing an error of reasoning? Is that rejection different from rejecting CH or NCH?

4 Logical Constants

In the last two chapters we looked at deviations from and extensions of standard logic. In Chapter 3 we discussed two *extensions*: the extension of FOL to $\mathrm{FOL}_=$, i.e. that of first-order predicate logic to first-order predicate logic with identity, and that of FOL to SOL, i.e. the extension to second-order logic. As we have seen, such extensions might significantly increase the expressive power of a logic and change its metalogical properties in dramatic ways.

The extension to $\mathrm{FOL}_=$ is motivated because it recognizes identity as a logical connective and doesn't treat it as a non-logical expression. But on the basis of which consideration can we decide which expressions should belong to the arsenal of logic? Why should identity belong to it but not the 'is larger than'-relation? What speaks against 'has a beard' as a logical constant? As we shall see in the second part of this chapter, the problem of how to demarcate the logical constants has no easy answer.

However, we will first need to attend to another, prior problem. Before we get to discuss whether we should add, say, identity to logical constants such as negation, conjunction, etc. we first need to understand better in virtue of what anything counts as *that* constant.

The Meaning of Logical Connectives

This problem arises when we consider our discussion of alternative, deviant logics in Chapter 2. As we have seen, some deviant logics are motivated by the conviction that certain inferences that are licensed by classical logic are not in fact valid. For example, dialetheists hold that a contradiction should not lead to "explosion" (a.k.a. *ex falso quodlibet*) because there may be true contradictions. But is a logic that considers expressions of the form $\ulcorner \phi \wedge \neg\phi \urcorner$ to be possibly true still a logic in which '¬' expresses negation?

Here is an argument (which Paoli (2003, 531) reconstructs from Slater (1995)) that paraconsistent "negations" are not negations:

1. Contradictories cannot be true together.
2. A sentence and its negation are contradictories.
3. If **L** is a paraconsistent logic, then, in the semantics for **L**, there are "inconsistent" valuations which assign both A and $\neg A$ a designated value, for some formula A.
4. If A and B both receive a designated value, under some valuation v, in the semantics for **L**, then A and B can be true together according to **L**.
5. In paraconsistent logics, A and $\neg A$ may not be contradictories (from (1), (3), (4)).
6. Thus, paraconsistent "negations" are not negations (from (2), (5)).

Is this right? If so, what does that imply for the nature of the conflict/disagreement between classical and deviant logics?

Quine's Thesis and his Critics

Quine (1970) has already argued that the dialetheist's system is not in fact representing negation, and that every change of logic is ultimately a change of subject:

> My view of this dialogue [between those who think that there can be true sentences of the form $A \wedge \neg A$ and those who do not] is that neither party knows what he is talking about. They think they are talking about negation, '[¬]', 'not'; but surely the notion ceased to be recognizable as negation when they took to regarding some conjunction of the form [$A \wedge \neg A$] as true, and stopped regarding such sentences as implying all others. Here, evidently, is the deviant logician's predicament: when he tries to deny the doctrine he only changes the subject. (Quine 1970, 81)

If we abstract from the particular elements of the example (the involved logics and the discussed connective), Quine's thesis is that in a difference between two logics, $\mathcal{L}_1$ and $\mathcal{L}_2$, a change of subject is involved. *Grosso modo*, two logics are different if and only if their collections of logically valid and invalid arguments are different. According to Quine, the change of subject consists in changing the way we understand the connectives, in changing their meaning. This, in turn, consists in changing the values that a given formula takes under certain conditions and also in modifying the entailments between certain sentences. So 'change of subject' means 'change in the meaning of the logical connectives' (although perhaps Quine would not put it that way given his skepticism about meaning), which in turn means 'change in the truth-values of certain formulas, presumably attributable to a change in the way of evaluating the connectives, and a change in the valid inferences between certain formulas'.[1] Thus, put in simpler terms, the meaning of a logical connective is determined by the truth-conditions of the sentences in which it appears. The fact that those truth-conditions validate certain theorems, namely, those of classical logic, is because those truth-conditions replicate the use of their everyday language counterparts as faithfully as possible.

This Quinean thesis is typically accepted. Consequently, deviant logicians argue for a change of logic and, thus, a change in the meaning of the connectives, on the basis of a purpose that requires this. An intuition common to several proponents of deviant logics is that if one has to modify the collection of classical theorems, this is because some theories we consider worthy for inclusion in our body of knowledge require different logics (see e.g. Priest, 2006; da Costa, 1974). As an example, if when studying quantum phenomena one has to use a logic that is not classical, this is because subatomic particles have properties the treatment of which in classical terms is inadequate or, at least, greatly hampers their study.

By accepting that there can be legitimate changes in the meaning of a connective, Quine's critics have accepted the thesis that a change of logic is a change of subject and thus that changing the subject changes the meaning of the connectives. Certainly there seem to be good reasons to accept this. For instance, possible-worlds semantics can be seen as making the difference between classical and intutionistic[2] negation explicit as follows

[1] Although, strictly speaking, one should speak about the truth-conditions of the propositions in which the connectives appear, for simplicity we will talk on occasion about "a connective's truth-conditions".
[2] See Chapter 10 for a brief explanation of intuitionistic logic.

(where $\mathfrak{v}$ is a valuation, w and w' are possible worlds and $\geq$ is an ordering or accessibility relation):

Classical negation: $\mathfrak{v}(\neg\phi, w) = \top$ if and only if $\mathfrak{v}(\phi, w) = \bot$.
Intuitionistic negation: $\mathfrak{v}(\neg\phi, w) = \top$ if and only if, for every $w' \geq w$, $\mathfrak{v}(\phi, w') = \bot$.

Another option – albeit a less explored one – is again to accept that there is a change of subject, but to argue that what changes is not the meaning of the connectives but instead our notion of deducibility (cf. Estrada-González, 2011). Again, this option is also compatible with Quine's thesis, since it merely shows that a modification in the collection of theorems and the meaning of the connectives, or in the deducibility relation, is called for by the requirements of some of our theories.

In a Quinean philosophy of logic and science, changes of inferential tools are considered to be global rather than merely local. It would not be possible, for instance, to maintain classical logic as a tool for every domain except that of, say, subatomic phenomena. Thus, Quine's thesis is that a change of logic is a change of subject, but also that this change has to be global.

We are not interested in the discussion of this stronger thesis, as it involves discussion of the nature of logic, the criteria for choosing between logics, etc. We are only interested in discussing here a weaker version of the Quinean thesis, namely that if $\mathcal{L}_1$ and $\mathcal{L}_2$ are different logics, then their connectives must have different meanings. In what follows, when we mention either 'the Quinean thesis' or 'Quine's thesis' we will be referring to the weaker version.[3] Also, when talking about *the problem of logical connectives* or *the debate about connectives* we will not refer to the perhaps more frequently discussed problem of finding out which are the logical constants, be it in a formal theory or for our everyday language, which will be dealt with later in this chapter. The problem for now isn't which or how many logical constants there are, but rather what determines the meaning of connectives, regardless of which and how many of the connectives actually count as logical constants.

[3] This thesis has been quite pervasive among many logicians' practices, more so than the stronger version. In the literature, this thesis has been explicitly defended by Priest (2006, Ch. 10) for the specific case of classical and intuitionistic logic.

Quine's thesis has been widely accepted even by those with different views on the meaning of the logical connectives. Among the few objectors to Quine's thesis one can find Hilary Putnam (1957, 1962, 1968), Adam Morton (1977) and, more recently, Greg Restall (2002), Jc Beall and Greg Restall (2001, 2006), Francesco Paoli (2003, 2007) and Stephen Read (2008).[4]

In his objection to Quine, Hilary Putnam (1962, 377) defends three theses:

(Put1) The connectives have a "nuclear" or "central" meaning (*core meaning*) which is independent of many of the theorems or valid arguments (or proofs, if one prefers proof-theoretic terms) in which they appear; that is, for Putnam, Quine does not adequately characterize the meaning of a connective.

(Put2) If one interprets 'change of meaning' as a modification on the global use of a connective, i.e. if by 'change of meaning' one means that the theorems (or valid arguments) associated with a connective do not match for two given logics, then (trivially) changing the logic is changing the meaning.

(Put3) If the thesis 'change of logic is a change of meaning' is taken to mean that a change of logic *only* involves a change of meaning of the connectives, then the thesis is false, as a change of logic affects the deducibility relation too.

Obviously, theses (Put2) and (Put3) matter just in case the meaning of the connectives is different from their core meaning. Thus, we will focus only on thesis (Put1) (which denies thesis (Put2)'s antecedent) and on Putnam's defence of it, which can be found in Putnam (1957).

Putnam traces a distinction, although rather implicitly, between the theorems (or the valid arguments) of a logic $\mathcal{L}$ in which a connective c appears and the properties such a connective must meet in order to be c (even if such properties could validate certain theorems or arguments). In the context of a discussion of the logic of quantum mechanics, Putnam

[4] Each of these authors is a "minimalist" in the sense introduced later in this chapter, although all are minimalists of very different sorts: while Putnam comes from an axiomatic tradition, Beall and Restall discuss the issue in model-theoretic terms and Paoli and Read do so in more contemporary proof-theoretic terms.

introduces the idea of an "operational meaning" for the logical connectives. Thus, the theorems in which disjunction appears would be one thing, and the features a connective needs to have in order to be considered a disjunction would be another. Leaving the discussion of quantum mechanics aside, and modifying Putnam's presentation slightly, the operational meaning of the connectives can be specified as follows. Suppose that there is a semantics in which there are at least two truth-values (say, *true* and *false*) and that these form a partial ordering. The operational meanings for the connectives would then be the following:

- the operational meaning of $\phi \wedge \psi$ consists in its truth-value being the infimum of both conjuncts' value;
- the operational meaning of $\phi \vee \psi$ consists in its truth-value being the supremum of both disjuncts' value;
- the operational meaning of $\neg\phi$ consists in the truth-value of the conjunction of $\neg p$ and p being equal to *false* and the disjunction of $\neg p$ and p being equal to *true*.[5]

The connectives of quantum and classical logic share this operational meaning; additionally, it is not difficult to show that the classical theorems can be characterized from this operational meaning once one restricts the collection of truth-values to two. Thus, according to Putnam, in the discussion of the relation between classical and quantum logics, the number of truth-values and the (in)validity of some formulas such as $(p \wedge (q \vee r)) \leftrightarrow ((p \wedge q) \vee (p \wedge r))$ are irrelevant for the specification of the meaning of the logical connectives. Taking only the above-mentioned operational meaning, the connectives of classical and quantum logic do not differ in their meanings.

Morton (1973) takes Putnam's position further by stating that there need not be a single collection of arguments (or theorems) that has to be shared by the connectives c, c' and c'' in order for them to be similar, for the three of them to be, for instance, disjunctions. Morton suggests that two

[5] This proposal by Putnam is not as general as one would like it to be: the operational definition for negation is of no use either for intuitionistic logic or for many inconistency tolerant logics. Of course, in the text under examination, Putnam only tries to show that there is an operational meaning common to classical and quantum logic, but in texts such as Putnam (1957), it is explicitly stated that the validity of $\phi \vee \neg\phi$ plays no role in the meaning of either disjunction or negation.

connectives c and c' are *similar* to each other if c and c' share a common collection I^* of valid arguments (theorems). However, *similarity* does not entail *sameness of meaning*, which is what it would take to argue against Quine's thesis. As the similarity relation is not transitive, a connective c'' could be similar to the connective c' if they share some collection I^{**} of valid arguments (theorems) even if I^* and I^{**} are not the same collection of valid arguments (theorems) and even if they are disjoint (and, therefore, even without c and c'' being similar). Putnam's proposal does not face this issue because in order for two connectives c and c' to be, say, a conjunction, they both have to validate a certain collection of theorems or arguments.

In the proof-theoretic framework, an *operational rule* for a connective c in a logic $\mathcal{L}$ tells us how c is used in the proofs of $\mathcal{L}$; this concerns introduction rules "to the left" or "to the right", as we encountered in Chapter 1.[6] Also, remember that no connectives appear in the *structural rules* of $\mathcal{L}$. There are logics that do not differ in their operational rules, but only in the structural ones, so it is at least dubious that the meaning of the logical connectives differs between *these* logics. However, those who endorse a proof-theoretical approach have not reached a consensus on whether the meaning of the connectives is only determined by the operational rules or if there is a contribution made by the structural rules too. Wansing (2000) assures us that, in practice, logicians often recognize that the connectives have two kinds of meaning: an "operational" meaning, determined exclusively by the operational rules for the connective, and a "global" one, in which the contribution of the structural rules is taken into consideration.[7] Regardless, the recognition of these two kinds of meaning is not of much use in refuting the Quinean thesis, as there is still at least one sense in which a change of logic *is* a change of subject, namely when – despite having the same operational meaning – the connectives of two logics differ in their global meaning.

Paoli (2003) is well aware of this difficulty, yet still holds that there can be a change of logic, and specifically a rivalry between logics, without there

[6] Gentzen (1969, 80) suggests that the meaning of a connective is specified by its corresponding (operational) introduction rule. This idea is further developed by Dummett (1978), Prawitz (1982) and Dǒsen (1989), among others.

[7] Notice that the operational meaning that is being talked about in the contemporary proof-theoretic context is not the same operational meaning that Putnam talks about.

being a change of meaning, as the connectives of different logics can have the same operational meaning, that is, the same operational rules (see Paoli, 2003, 539). However, this proposal suffers from a problem similar to that of Morton's: that the operational rules are the same for a connective in two different logics says at most that both connectives belong to the same kind (negation, for example). In fact, one can even say that the sharing of the operational rules by two connectives entails that both connectives share some common meaning. However, this does not warrant that those rules are sufficient to specify the *full* meaning of the connectives, which leaves room for there being a change in the full meaning of the connectives among different logics if the latter differ in their structural rules. An argument to discard the contribution of structural rules is needed.

Maximalisms and Minimalisms

In the more recent debate concerning the meaning of connectives, one can identify two camps. One is that of model-theorists; their main thesis is that the meaning of a logical connective is determined by its contribution to the satisfiability conditions of formulas that contain them. The other camp is that of proof-theorists, according to which the meaning of a logical connective is specified by the connective's contribution to the inferential role of formulas that contain it.

Thus, for example, according to proof-theorists, the meaning of conjunction consists in that from $\phi \wedge \psi$ one can infer ϕ as well as ψ (as in a conjunction elimination rule), and that from ϕ and ψ one may conclude $\phi \wedge \psi$ (as in a conjunction introduction rule). That would be the contribution of conjunction to the inferential role of the formulas in which it appears and, hence, that would be its meaning. Both the model-theoretical and the proof-theoretical versions of the Quinean thesis are usually accepted, that is, for model-theorists changing the logic changes the satisfiability conditions for the connectives, and for proof-theorists changing the logic changes the collection of valid arguments via a change in the rules of use for the connectives.

Typically, both model-theorists and proof-theorists are, to borrow a useful expression from Restall (2002), "maximalists". According to *semantic maximalism*, every semantic feature related to a connective c contributes to its meaning. It is typically held that the meaning of a connective is determined by its satisfiability conditions, which in the case of many zeroth-order logics are usually expressed in truth-tables. But truth-tables for typographically identical connectives may differ radically, which gives Quine's thesis more plausibility. Model theory seems to be tied then to semantic maximalism. *Model-theoretic maximalism* is the thesis that every element involved in determining the satisfiability conditions for formulas in which *c* is the main connective contributes to its meaning – those elements involve the details on the structure of values, indices, domains, accessibility relations, etc.

Maximalism also comes in proof-theoretic terms. For the sake of definiteness, we will stick to sequent calculi, as we have done in the rest of the book. *Proof-theoretic maximalism* is then the thesis that the meaning of any connective *c* is determined both by the L- and R-rules for *c* as well as the structural rules. The meaning of a connective is its "global meaning", to employ the useful terminology of Paoli (2003).

According to *semantic minimalism*, only *some* semantic features related to a connective determine its meaning. The obvious problem is picking the ones that do. As we have seen in the case of Putnam, only some theorems in which *c* appears contribute to determine its meaning. In Putnam's case study, the answer as to what are the meaning-determining elements is that they are the theorems of a certain quantum logic, as they constitute the common core of both this quantum logic and classical logic. The problem is that this strategy is not so easily extended to cases in which one wishes to consider more logics, let alone *all* logics.

Proof-theoretic minimalism is the thesis that only some rules contribute to determine the meaning of a connective *c*. Typically, just the operational rules are taken to play such a role.[8] This choice both motivates and is motivated by

Došen's Principle If two logics presented proof-theoretically are different, their difference lies in elements of the proof theory other than the operational rules for connectives (see Došen, 1989).

8 Nonetheless, some authors think that *tonk* refutes the idea that operational rules alone can determine the meaning of connectives. We will say something more about *tonk* below.

Now, even if one agrees that it is just the operational rules that contribute to determining the meaning of connectives, this is still a generalization. Do both L- and R-rules play such a semantic role, or rather just the R-rules (as Gentzen thought) or only the L-rules? Perhaps the kind of rule that makes the semantic contribution varies depending on the connective, as each rule extracts information in different ways, which might suit some connectives but not others?

Besides these choices, a proof-theoretic minimalism can be asked to satisfy certain desiderata like *separation* (characterize *c* without using other connectives); *weak symmetry* (every rule is either an L- or an R-rule); *weak explicitness* (*c* is present only in the bottom sequent), etc. Again, leaving aside the details of specific proposals, one nice feature of proof-theoretic minimalism is that it can provide useful distinctions between kinds of meanings: a local one, when only operational rules make a semantic contribution, and a global one, when structural rules also make a contribution. This in turn provides the means to refute Quine's thesis: if the only, or at least the main, meaning of a connective is the logical one, there are plenty of cases of different logics that share the operational rules and only differ in structural ones, so there is a difference between logics without meaning-variance, after all.

After the initial proposal of Restall (2002), there has been little work on model-theoretic minimalism. What is a satisfiability condition? It is usually considered to be a *specific* function from constructions on truth-values to truth-values. By "specific" we mean that the details of the domain and codomain of the function are fully specified: the exact number of truth-values (not only assuming that there is at least one), the exact ordering between them, the exact number and nature of indices and accessibility relations, and so on. But that sounds more like maximalism, not minimalism.

Consider a Došen's Principle for model-theoretic semantics: If two logics are different, their difference lies in elements of the model theory other than the satisfiability conditions. In order to give content to model-theoretic minimalism, it should be possible to distinguish between a kind of rule of association, an "operational (or "local") satisfiability condition", and a particular map in which all the parameters of the rule of association are properly filled in, a "global satisfiability condition". So one would have at least the same nice features as in proof-theoretic minimalism but probably also the same problems, *mutatis mutandis*. Work in that direction can be found in Estrada-González (2011).

Even though attractive, the minimalist thesis, whether in its model-theoretic or proof-theoretic version, still has to overcome many difficulties both in its enunciation as well as in its formulation of a more or less complete proposal. So far there is no conclusive argument to decide which elements are relevant for specifying the meaning of the connectives and which are not.

Are Disputes about Logic Merely Verbal?

So far we went along with the idea that if the connectives in two logics are governed by distinct rules, this would involve a change of meaning. The way that Quine thought about it, as we have seen, implies that whenever two logicians disagree about the rules that should properly govern a connective, they in fact talk past one another, since they talk about different connectives. The proposals that we have looked at so far, which could avoid this last implication, were such that they tried to find a set of rules of some kind that would remain stable between the disagreeing logicians, such that due to these stable rules the logicians could still be said to talk about the same connective. Perhaps this isn't the only way in which stability of meaning could be guaranteed between interlocutors.

Let us take an example: let's consider Catharina and Stephen, two logicians. Stephen endorses some relevant logic, while Catharina endorses classical logic. Let us consider their dispute under three different conditions. In the first condition, Stephen and Catharina are both native speakers of English. Moreover, they are both disposed to reason according to, say, classical logic. In other words, their *logica utens* is the same, but their *logica docens* differs. Thus, when it comes to metalogical arguments, which are carried out in English, they both reason classically. However, Stephen arrived at the view that Catharina's classical logic is wrong. Do Catharina and Stephen now talk past one another, when they disagree about the rules that govern conjunction, negation or the conditional?

In this first case, it seems that they don't. At least one of them, probably Stephen, is simply mistaken about what the correct *logica docens* is. Perhaps within the proposed theoretical frameworks, the formal counterparts of the logical connectives have different "meaning", but at least one of them is just an inadequate representation of the actual logical constants.

Let us change the case a bit and consider a situation in which Stephen's *logica utens* conforms to his *logica docens*. After years of studying relevant logic, he developed a disposition to reason in that logic too. Now his metalogical proofs are also governed by the rules of relevant logic in the sense that the inferences he in fact draws are in accordance with these rules. In this second case we might think that there is now a better case to be made for Catharina and Stephen to talk past one another when discussing matters of logic.

However, Timothy Williamson (2007) has argued for a view on which this second case is not actually unlike the first. Catharina and Stephen should both still count as speakers of the same language. They both learned English as their first language, and both clearly count as competent speakers. When they talk about the rules that govern the logical connectives, they talk about the rules that govern 'and', 'or' and 'it's not the case that' etc. The meaning of these words in English is determined by the use of these connectives in the linguistic community as a whole. Catharina happens to be attuned to that meaning, Stephen's reasoning is out of step with it, but that just means that he sometimes makes *mistakes*. When Catharina and Stephen use these words in their metalanguage, they nevertheless use these words with the same meaning, the meaning these words have in English. Their individual quirks cannot change the meaning of these words. This follows from a relatively uncontroversial social externalism about meaning on which the meaning of expressions in a public language is determined by the linguistic community, and on which individual members of the community count as members on the basis of sufficiently successful linguistic interaction with that community.

Let us consider a third case, on which Catharina and Stephen do not actually speak the same language. Let's assume that Catharina speaks British English, while Stephen speaks Australian English. Let us also assume further that the linguistic community in Australia, after decades of exposure to deviant logics, has developed a disposition to use the English words for 'and', 'or', 'it's not the case that', etc. as if their meaning was governed by the rules of relevance logic. On the externalist picture above, Catharina and Stephen might now count as speaking two different languages with the logical connectives in each having different meanings (and thus, perhaps, should count as different connectives, depending on what stance one takes on the topics discussed in the previous sections).

However, the fact that some homophonic words in British English and Australian English happen to have different meanings does not imply that therefore Catharina and Stephen must be talking past one another. First of all, they might be aware of the fact that these words in their respective languages have different meaning and thus might be perfectly able to avoid merely verbal disputes, even if both keep speaking their language. In this case, it would then be a question of what they take themselves to be discussing – perhaps they discuss the rules as they apply to the connective in one of the languages, or perhaps they discuss the question of which rules should apply, or they discuss the structure of *logica ens*, assuming that this is independent of what language people happen to speak.

Let us consider that last option a bit further. As we shall see in the next chapter, logical realists hold the view that logic studies a mind-independent subject matter. Such a subject matter could then well be independent of the inferential dispositions of any linguistic community. Catharina and Stephen might intend to study that. In fact, regardless of whether the words 'and', 'or', 'it's not the case that', etc. have different meanings in their native languages, they might be using these words as technical terms with the stipulation that their meaning should "carve nature at the joints". If Ted Sider (2011) is right, then the structure of the world is such that it provides meaning candidates for the logical connectives.

> First question: do any logical concepts carve at the joints? There is a powerful argument that at least some of them do. The best guide to joint-carving is a Quinean criterion of ideological commitment: it is (defeasibly) reasonable to regard indispensable ideology as carving at the joints. But we cannot get by without logical notions in our fundamental theories. In particular, since Frege it has become clear that the notions of first-order predicate logic are indispensable in serious foundational theorizing. (Sider, 2011, 216)

Thus, because logic is indispensable we have reason to believe that the concepts of logic carve at the joints. But then there is an objective structure that Catharina and Stephen can intend to be the meaning of their expressions, just as we might intend natural kind terms like 'water' to be expressing as a concept whatever it is that in fact is water. The meaning of logical connectives then isn't in the head, and not in anyone's inferential dispositions either, be it at the level of individuals or at that of linguistic communities.

This is a "voluntaristic" model of how Catharina and Stephen might come to use the logical connectives with the same meaning, even if they are both members of different linguistic communities, each with a different *logica utens*. It is "voluntaristic" because it depends on Catharina and Stephen both intending to use logical connectives in this way. This is perhaps an unrealistic assumption. While there might be empirical evidence that we use natural kind terms with a deferential disposition (deferring to the actual nature of, for example, water for the determination of the meaning of 'water'), it is not plausible to hold that we have the same disposition when it comes to logical constants.

According to Sider (2011), the same result may, however, be obtained from a non-voluntaristic picture. According to "reference magnetism", expressions for which there are joint-carving meaning candidates will also have these as meanings, regardless of whether speakers also intend to use these expressions with these meanings. For example, let's assume that the classical meaning of the logical connectives corresponds to the world's structure (as understood by Sider), then the logical connectives of British English and of Australian English (in our hypothetical example) will both have that same meaning, even though the use of the logical connections in Australian English seems to be somewhat different from that in the British English linguistic community. However, the fact that the world has the structure that corresponds to British English will then trump the dispositions of the speakers of Australian English.

Reference magnetism of this trumping kind is perhaps not a terribly plausible position (cf. Schwarz (2014) for a critique of reference magnetism). Sider (and before him, David Lewis (1984)) defend the view as a solution to Putnam's challenge of reference skepticism.[9] We cannot discuss

[9] Hilary Putnam (1980) argues that we can find for any global theory infinitely many interpretations that satisfy that theory. This is Putnam's model theoretic argument against "global descriptivism". The theoretical role of an expression, even in a global theory, doesn't fix the extension for that expression. If there is one domain that will satisfy the theory, then every domain of objects of the same cardinality will do the same. Moreover, Putnam added, any constraint other than truth (i.e. satisfaction of

the matter in any detail here. However, on this view the deviant logician could not talk past the classical logician; they would in fact use the logical connectives with the same meaning, and that would be guaranteed by the objective structure of the world.

This shows that Quine's thesis, namely that the meaning of logical connectives varies between different logics, does not also entail the conclusion that Quine is drawing from it, namely that the proposal of an alternative *logica docens* means a change in subject, and that logicians of different camps must be talking at cross-purposes.

The Problem of Demarcation

Logic is usually thought to concern itself only with features that sentences and arguments possess in virtue of their logical structures or forms. The logical form of a sentence or argument is determined by its syntactic or semantic structure and by the placement of certain expressions called 'logical constants'. Thus, in order to determine which sentences are logically valid and which sentences logically true, we must tell apart the "logical constants" of a language from its non-logical expressions in a principled way. How to make such a demarcation is *the problem of logical constants*.

As we have observed in the previous chapter, there can be considerable disagreement over which expressions should be considered logical constants. We have seen that sometimes – for example in the case of the identity sign – the decision might be relatively easy: adding identity to the stock of logical constants enriches the expressive power of the language, but the metalogical properties of the logic stay otherwise largely the same, and we don't need to make any larger changes to model or proof theory in order to accommodate identity as a new logical constant.

As we also observed in Chapter 4, this is often not the case though. Adding higher-order quantifiers, constants from set theory or mereology, modal or tense operators, etc. will considerably change the logic so

the theory by a model) would be just another piece of theory, hence not a way out of the embarrassment of riches. Lewis (1984) suggests solving the problem from the other direction; instead of having the theory fix the extension of the expressions in the world, the world fixes the meaning of the expressions by being structured into eligible candidate meanings. It is controversial to what extent Lewis does in fact present this idea as part of his own view on meaning (cf. Schwarz, 2014).

enriched. Moreover, unlike in the cases of conjunction, negation, disjunction, etc. we don't have strong or very widely shared intuitions about whether these expressions should count as logical constants. For these cases it would be good to have a principled demarcation criterion that told us where to draw the line between the logical and the non-logical expressions.

Before turning to specific proposals on how to draw the line between logical constants and other expressions, we should say a bit more about the status of the problem and its motivation. Following MacFarlane (2015), we can distinguish four general stances towards the problem of logical constants: those of the Demarcaters, the Debunkers, the Relativists and the Deflaters.

According to the *Demarcaters,* logic is the study of the properties that arguments or reasonings have in virtue of the *logical form* of premises and conclusions. However, logical form is in turn determined by the logical constants that occur in these premises and conclusions. Thus, the subject matter of logic is, in fact, determined by where we draw the line between logical and non-logical expressions.

For a Demarcater the question of where we draw that line is therefore a central issue in the philosophy of logic, and a theory that would enable us to draw a principled distinction between logical and non-logical expressions would also be expected to provide us with substantial insight into the nature of logic. Typically Demarcaters are *optimists*, for they think the problem is not only genuine and important, but solvable by means of some theory, whether mathematical, semantic, epistemic or any other kind.

In contrast, *Debunkers* do not think that the subject matter of logic is determined by the choice of logical constants. Logic instead studies logical validity *simpliciter*. Logicians study that relation by classifying arguments by their form, but that is a *tool* that logicians use and not the proper subject matter of logic. Consequently, the demarcation problem is really just a pseudoproblem.[10]

[10] These characterizations leave room for what might be called "Optimist Debunkers". Their view would be that, even if logical forms and logical constants do not form part of the subject matter of logic – such that the problem of logical constants does not have the centrality Demarcaters give to it – a precise characterization of logical constancy is nevertheless possible.

As MacFarlane (2015) explains, one can illustrate the difference between Demarcaters and Debunkers by looking at how they think about the value of formal counterexamples. For a Demarcater it is possible to show that a natural language argument is *invalid* by providing a formal counterexample: another argument that has the same logical form, but true premises and a false conclusion.

However, as we already noted in the previous chapter, one and the same natural language argument can be moulded into different logical forms. Thus, for Debunkers, counterexamples never show anything about a particular argument. All they show is that a certain argument *form* is invalid, that is, that it has invalid instances. But every argument can be seen as an instance of

$$\frac{\phi}{\psi}$$

and this hardly entails that no argument is logically valid. A formal counterexample just shows something about *that* form of the argument, but not about the actual premises and conclusions of arguments. Thus, an argument like

1. Jim's car is orange.
2. Jim's car is coloured.

could count as logically valid for a Debunker if it is impossible to find a case in which *that* premise is true but *that* conclusion false.

The Demarcater, on the other hand, will reply that a genuine counterexample to the formal validity of an argument would have to exhibit its *full* logical structure. Thus the Demarcaters' use of counterexamples to demonstrate the formal invalidity of arguments presupposes a principled way of discerning the full logical structure of an argument, and hence of distinguishing logical constants from non-logical constants.

Furthermore, the Demarcaters can present the Debunkers with a dilemma: either endorse a (metaphysical or epistemic) conception of analyticity that would show the argument above about Jim's car to be *logically* valid, but the following argument to be merely non-logically valid

1. The particle a has mass.
2. Therefore, the particle a does not travel as fast as light.

(first horn of the dilemma); or logic must be a universal science that can explain all cases in which an argument's premises somehow guarantee the truth of that argument's conclusion (second horn of the dilemma).

The first horn of that dilemma seems unacceptable to many contemporary philosophers who think that Quine has shown that the analytic/synthetic distinction is untenable. We will discuss Quine's view in Chapter 6. The second horn doesn't seem to be any better, as long as the relevant notion of necessity isn't further specified. The inference from a particle's mass to that particle's possible speed seems clearly outside the realm of logic, but perhaps there is a way to distinguish a relevant notion of *metaphysical* necessity (which grounds logical consequence) from other notions of necessity (such as, for example, physical necessity). We will come back to this idea in Chapters 5 and 10.

Relativists have a middle position between Demarcaters and Debunkers. They agree with Demarcaters that logical consequence is *formal* consequence and hence requires a distinction between logical and non-logical expressions. At the same time, however, they agree with the Debunkers that such a distinction can't be *absolute*. Instead, logical consequence needs to be *relativized* to certain choices of logical expressions. For each such set, there is a corresponding notion of logical consequence, and each of these may have its application. In the extreme case, when we consider all expressions of a language to be logical, logical consequence collapses into material implication. Relativism about logical consequence is thus already one form of *logical pluralism*: the idea that there can be more than one correct logic. We will discuss this topic in more detail in Chapter 7.

The last stance towards the demarcation problem that MacFarlane (2015) distinguishes is that of the *Deflater*. We said above that Demarcaters believe that there is a principled distinction between logical and non-logical expressions, and that this distinction can be made on the basis of a substantial theory which should be illuminating about the nature of logic. Deflaters doubt the latter. They agree that a distinction can be made and that the distinction itself is important, but they doubt that there is a substantial and illuminating story to be told in virtue of what the distinction holds. Just like there (arguably) is no illuminating story to be told about which activities are "games" and which aren't, there is also no such illuminating story about which expressions are logical and which aren't.

Thus, Deflaters agree with Demarcaters that there is a real distinction between logical and non-logical constants, and between formally and materially valid arguments, both of which are important for understanding the nature of logic. But Deflaters reject the Demarcaters' project of finding precise and illuminating necessary and sufficient conditions for logical constancy for all languages, or for an all-purpose single language.

As MacFarlane (2015) observes, the proof should be "in the pudding": can Demarcaters actually provide illuminating necessary and sufficient conditions for logical constancy? Let's look at some proposed demarcations.

Model-theoretic Demarcaters

The most promising approach to demarcate the logical constants from other expressions is the "permutation invariance" approach (cf. Sher, 1991; Tarski, 1986). Logic is supposed to be "topic neutral", logical connectives aren't supposed to be about anything. In model-theoretic terms this could be cashed out by requiring that the extension of a logical constant should be independent of what the domain is like. To check whether an extension is independent of what the domain is like, we can just see whether the extension stays the same if objects in the domain "swap places" so to speak, if we consider *permutations* of the domain. In a permutation, objects from a domain are mapped onto other objects in that domain. We can then define *transformations* relative to permutations. A transformation tells us how types of the hierarchy depend on the permutation. Take as an example the domain of all coloured things and two proper subsets of that domain: the set of all the red things and all the green things. If we permutate that domain, then it might be that the permutation maps an object that happens to be in the set of the red things onto another object that happens to be in the set of green things, say. The transformation of our original set of red things has now a different extension. There is now an object in that set that wasn't in the original set prior to the permutation, so to speak. On the other hand, if we take the identity relation as a set of ordered pairs, then that set will remain the same, regardless of how we permutate the objects in the domain. The identity sign, which has that set as its extension, is thus "permutation invariant".

One needs to play around with the details of the approach a bit, in order to get the intended result, but the permutation-invariance criterion delivers:

> The monadic predicates "is a thing" (which applies to everything) and "is not anything" (which applies to nothing), the identity predicate, the truth-functional connectives, and the standard existential and universal quantifiers all pass the test. So do the standard first-order binary quantifiers like "most" and "the". Indeed, because cardinality is permutation-invariant, every cardinality quantifier is included, including "there are infinitely many", "there are uncountably many", and others that are not first-order definable. Moreover, the second-order quantifiers count as logical (at least on the standard semantics, in which they range over arbitrary subsets of the domain), as do all higher-order quantifiers. On the other hand, all proper names are excluded, as are the predicates "red", "horse", "is a successor of", and "is a member of", as well as the quantifiers "some dogs" and "exactly two natural numbers". So the invariance criterion seems to accord at least partially with common intuitions about logicality or topic neutrality, and with our logical practice. (MacFarlane, 2015)

This proposal is thus philosophically motivated, mathematically precise, many of its results accord with common practice of logicians, and it decides some borderline cases. As Denis Bonnay (2014) argues, it also satisfies further desiderata that distinguish it from alternative approaches: it is *pure*, *local* and *intrinsic*.

The invariance criterion is *pure* insofar as it only considers semantic properties (rather than grammatical or proof-theoretic ones) for the characterization of the available interpretations for the logical constants. It's *local* since the criteria apply to the semantic properties of the expression in question rather than to the expression as a part of a larger system (as approaches do which involve meta-logical properties for the demarcation of the logical constants). Finally, the invariance criterion is *intrinsic*, because the logical notions are not contrasted or compared to other notions (Bonnay, 2014, 55).

Problems for the Invariance Conception

The invariance criterion considers permutations of a given domain. That way, the existential and the universal quantifier of classical logic come out as being permutation invariant. In fact, the invariance criterion requires invariance under bijections, which preserve the cardinalities of domains. But then all quantifiers with finite ('there are exactly seven') and even infinite cardinalities ('there are uncountably many') pass the invariance test. Intuitively, we might not want to consider these quantifiers to be logical (they are not even definably in FOL). But perhaps this is insisting too much on a general and principled boundary between logic and mathematics. Gila Sher, one of the most prominent advocates of the invariance approach, welcomes this result as an indication that logic and mathematics are one and the same:

> The formalist account of logic, with its Invariance-under-Isomorphism criterion of logicality, offers an explanation of the relation between logic and mathematics [...]. Mathematics, in this account, builds a theory of formal structure, and logic provides a method of inference based on this theory. I will call the new approach "mathematicism". (Sher, 2008, 318)

Thus, whether you see the logicality of these quantifiers as a problem depends very much on which view you hold concerning the relationship between logic and mathematics.

It also might be seen as a problem for the invariance account that constants are only considered as their extensions, rather than their meanings. Assume that it is both metaphysically and epistemically necessary, as well as analytically true, that there are no male widows. Then, a connective '$\bowtie$', defined by

$$\mathfrak{v}(\bowtie A) = \top \text{ if and only if } \mathfrak{v}(A) = \bot \text{ and there are no male widows}$$

would count as a logical constant (Gomez-Torrente, 2002, 21), since its extension is the same as that of '$\neg$'. Moreover, given that the nonexistence of male widows is an analytical truth, as well as an epistemic and metaphysical necessity, '$\bowtie$' threatens any attempt of supplementing the invariantist approach with a modal or a semantic requirement, like "permutation-invariance across all metaphysically possible worlds", "permutation-invariance across all epistemically possible worlds" (as in McCarthy, 1987) or "permutation-invariance implied by the meaning of the connective" (as in McGee, 1996), even without considering that they would be introducing notions that are not clearer than that of a logical constant. Appealing to a logical modality – requiring, for example, that logical constants have permutation-invariant extensions as a matter of logical necessity – faces the same problem of explicating the notion of a logical constant in terms of a primitive notion of logical necessity that is not clearer than that of a logical constant. On the other hand, if we don't treat *logical necessity* as a primitive notion, we run the risk of circularity: to define logic, one needs to define logical constants; to define logical constants, one needs logical necessity, and to define this one needs to know what makes something logical, and to know that, one needs to know what logic is.

In light of these difficulties, MacFarlane (2005) concludes that the permutation-invariance criterion should probably best be understood as merely providing a necessary condition for logical constancy, due to the fact that it operates only on the level of reference (not on the level of sense).

Proof-theoretic Demarcaters

One response to the last problem discussed could be to try to look at how one grasps the meaning of logical connectives and then to define as the logical constants those expressions the meaning of which is grasped in this special way. A more concrete suggestion could be to consider the logical constants to be exactly those expressions the meaning of which is fully determined by some rules in an appropriate proof-theoretic framework.

There are at least three problems with this approach. One problem is that not any set of rules will do for determining a logical constant. Famously, Arthur Prior (1960) devised the example of *tonk* in order to show that not every set of (introduction and elimination) rules fixes a meaning.

Prior's rules

$$\frac{\phi}{\phi \; tonk \; \psi} \; (\textit{tonk}\text{-introduction}) \qquad\qquad \frac{\phi \; tonk \; \psi}{\phi} \; (\textit{tonk}\text{-elimination})$$

permit inferring anything from anything. The rules combine the introduction rule of disjunction with the elimination rule of conjunction. The result of adding this constant to our standard arsenal (and associated rules) leads to disaster.

However, the fact that some combinations of rules do not determine the meaning of a constant does not establish that the whole approach is mistaken. If it could be shown that the right constants get determined if the rules observe certain extra conditions, then this approach could well work.[11] We discussed the idea that the meaning of a logical constant is provided by introduction and/or elimination rules already in the first half of this chapter.[12]

A second problem is that, even if one buys the general idea that the meaning of an expression is provided by rules of usage, it's not clear that the rules we have so far considered are sufficient for grasping the meaning of any expression.

Authors such as Dummett (1991), Sainsbury (1991), Gomez-Torrente (2002) have argued that the typical proof-theoretic rules for at least some logical constants do not exhaust all aspects of the use of these constants that must be mastered if one is to understand them. For example, they suggest that inductive reasoning from instances of a generalization to the generalization itself are partially constitutive of the meaning of 'all' and that it is not clear whether this part is, or should or could be, captured by rules for 'all'. Perhaps the situation can be ameliorated if one considers more complex sets of rules instead of pairs in determinate proof-theoretic presentations of logic, but we will not explore this suggestion here.

Alternatively, one could limit the project from a determination of *sense* to a determination of reference (or semantic value). This is the

[11] Moreover, it is not clear that from the fact that a logical connective permits inferring everything, as in the case of *tonk*, it follows that such a connective does not have meaning. It might trivialize the logic, in the sense that every connective has the same meaning as any other, but an additional argument is needed to the effect that when everything holds, or means the same, in a language, then the language is meaningless, and that what caused the trivialization is meaningless. See Priest (2006, chapters 1 and 3) for further discussion. Additionally, we mentioned that the context of deducibility also plays a role in the trivialization of the logic. If the underlying notion of logical consequence happens to be non-reflexive or non-transitive, *tonk*-like connectives need not trivialize the logic. We will discuss the general motivations for non-reflexive and non-transitive logic in Chapter 9; for applications of these to *tonk*-related issues, see Cook (2005), Ripley (2015), Fjellstad (2015). For a *tonk*-like connective which trivializes even under non-transitive logical consequence, see Wansing (2006).

[12] For other attempts to articulate conditions under which introduction and elimination rules do fix a meaning, see Belnap (1962), Hacking (1979), Hodes (2004).

approach that is followed by Hacking (1979). As Gomez-Torrente (2002) shows, this project then faces difficulties very similar to the “male-widow”-problem we discussed above for the permutation invariance approach:

> Consider the first-order quantifier ‘not for all not . . . , if all are not male widows, and for all not . . . , if not all are not male widows’. It has the same extension as the usual first-order existential quantifier. If it looks too long, [. . .] abbreviate it with ‘$\underline{\exists}$’ and think of it as a non-complex predicate with the already explained meaning. [. . .] [‘$\underline{\exists}$’] is a logical constant according to Hacking’s criterion. This is so because the *same* typical Gentzenian operational rules for the usual first-order existential quantifier hold for ‘$\underline{\exists}$’. [. . .] Thus, if Hacking’s procedure for “reading off” the semantics from the rules is right, then ‘$\exists$’ and ‘$\underline{\exists}$’ have the same extension. Their denotation is in both cases "fixed" by the rules, in the sense of Hacking. (Gomez-Torrente, 2002, 29)

However, ‘$\underline{\exists}$’ is not a logical constant, although Hacking’s criterion rules it in. Thus, this approach also is extensionally inadequate.

Pragmatist Deflaters

The principles guiding logicians (both implicitly and explicitly) in cataloguing certain expressions as logical have been essentially pragmatic principles with a considerably vague content; just consider the topic-neutrality criterion. As we have seen, logicians and philosophers of logic have often tried to offer philosophically richer characterizations of the notion of a logical expression than the ones suggested by the vague pragmatic principles. These attempts usually characterize the notion of a logical expression (or of the logical expressions to be found in a restricted set of expressions) in terms of alleged semantic, epistemic or mathematical peculiarities of the logical expressions. According to pragmatists, those attempts will fail, as they have failed, since it is impossible to model closely, let alone exactly, in semantic, metaphysical, epistemic or mathematical terms the vague and pragmatic notion of a logical expression.

Traditionally, logic has been seen as a discipline that deals with the most general traits of reasoning, features of reasoning that apply in all argumentative fields. It is therefore reasonable to think that one principle underlying choices of expressions as logical is that logic must deal with arguments which are correct in virtue of the properties of expressions employed in reasoning in general, expressions not specific to any argumentative field but common to all. This is a leading idea behind most of the Demarcaters' proposals. However, it seems that their principle suggests only a necessary condition for the logicality of an expressions, and the continued failure of the Demarcaters' attempts seems to confirm that. According to Gomez-Torrente (2002), logicians probably require at least implicitly that logical expressions play a relevant role in reasoning in general, or that their study be useful in the resolution of particularly significant problems in reasoning. If the pragmatist is right about this, there is a great deal of vagueness and complexity in the intuitive concept of a logical expression, and there is little chance to avoid that when formalizing such expressions as logical constants in a formal language. The pragmatic principles underlying the use of this concept leave a lot of space for divergent conceptions, and for incompatible ideas on which expressions will eventually count as logical. Nevertheless, as we have seen when discussing the Deflaters' stance, this does not mean that the question of which expressions are logical is arbitrary, since the mentioned principles are not compatible with just any idea about which expressions are logical. Vagueness does not imply arbitrariness.

Questions

1. Can you think of reasons why only the R-rules (or, perhaps, only the L-rules) of the operational rules for connectives should determine the meaning of a connective?
2. In the thought experiment involving Catharina and Stephen, we assume in the second version of the thought experiment that Stephen's *logica utens* is in line with his *logica docens*. The way he in fact reasons (also on reflection) follows the rules of the logical theory he endorses. Let's assume, however, that most speakers of English are like Catharina and use and endorse a different logic. Is there room to argue – *pace* Williamson – that Stephen speaks a different language?
3. How damaging is it that the permutation invariance criterion does not discriminate between coextensional constants?

5 The Metaphysics of Logic

When we talk about logical truths it seems to make sense to ask what it is that makes these truths true. This question can be understood in different ways, and we will need to disambiguate these in this chapter. Without entering into the territory of theories of truth,[1] one account of the truth of, for example, 'Snow is white' is that this sentence is true because snow is actually white. Should we say the same about 'Snow is white or snow is not white'? Is the latter true because snow is actually either white or not white? In other words, is that sentence true because of snow and its colour? This seems to be missing the point. Why 'Snow is white or snow is not white' is true is something you learn in a logic class. Moreover, there you learn to determine that this sentence is true without investigating snow or its colour.

This leads us to another way to understand the topic of this chapter: what is the subject matter that we investigate when we study logic? Is the subject matter of logic a set of very general facts in the world (for example, the general fact that snow's being white or not white is an instance of), or is logic the study of general facts about our minds? (After all, logic is said to lay down the "laws of thought".) Or is logic just the study of linguistic conventions (for example, how to use 'or' and 'not') – perhaps conventions we introduce and shape in part by the very study of logic?

Realism

The default position with respect to pretty much every area of human discourse that allows for the application of a truth predicate to the statements

[1] Theories of truth are not discussed in this introduction. For a good introduction to different conceptions of truth, see Künne (2003). As we shall see in a minute, the topic of this chapter can also be specified without mentioning truth.

made in that discourse is *realism*. In this chapter we will follow Michael Resnik (2000) and characterize realism with the following two necessary conditions:

> Logical realism is committed to at least two theses: First, there is a fact of the matter of whether something is a logical truth, a logical inconsistency or logically implies something else. [. . .] Second, that such facts (or the truth-values of such claims) are independent of us, our psychological make-up, our linguistic conventions and inferential practices. (Resnik, 2000, 181)

The first of these two necessary conditions, let us call it the "cognitivism condition", has two aspects. One is the idea that claims about logical truth, logical inconsistency, logical implication, etc. are *true* or *false*. The second aspect is that these claims have *representational content*; the claims are true or false because they correctly or incorrectly reflect certain facts.

The second of the two necessary conditions, let's call it the "objectivity condition", then states that whatever these facts are, they are independent of us ("our psychological make-up, our linguistic conventions and inferential practices"). But if these aren't the relevant facts, which other facts can plausibly matter for claims about logical truth, implication, etc.? There are at least two somewhat popular candidates.

Two Examples of Realism

Platonist realists hold that the facts that are correctly described by a true logical theory are facts involving certain *abstract objects*. If we are realists about, for example, propositions as certain abstract objects that are the contents of thoughts and sentences, then it might be relatively natural to hold that the implication relations between them and whether or not they are logically true are themselves facts that are just as objective and independent of us as the existence of these abstract objects themselves is. After all, these logical properties and relations will plausibly hold because of the intrinsic properties of those abstract objects. On this view, the logical facts obtain in a certain realm of abstract objects, a "third realm" independent of the mind but also independent of the physical world.

Structuralist realists, on the other hand, hold that the facts that the true logical theory correctly describes are very general metaphysical structures of this world.[2] For present purposes, the main difference between these two realist views seems to be how differently they construe logical facts from other facts. Whereas the platonist realist assumes a dichotomy in the realm of facts, the structuralist realist can consider the logical facts to be in the same realm of facts as all other facts (and isn't committed to the existence of abstracta).

Realists of the platonic kind will include philosophers like Edmund Husserl (1859–1938). In the first edition of the *Logical Investigations* (1900–1901), Husserl holds the view that it is features of independently existing abstract entities (*Bedeutungen*) that account for the truth of logical laws (cf. Lapointe 2014, 190).[3] More recent platonist realists include Jerrold Katz, who argues for logical realism (amongst other realisms) as realism about abstracta (Katz, 1998), and Penelope Rush (2014), who defends a Husserl-inspired variant of logical realism.

The structuralist realist, on the other hand, holds that logical theory describes very general (metaphysical) structures of the world. Logical theory is thus a very general metaphysical theory. There are several ways how one might arrive at such view. One way is to observe that logic is not – as it is often held to be – metaphysically neutral, but actually makes metaphysical claims. As we have seen in Chapter 2, many objections against standard logic are metaphysically motivated:

> A natural metametaphysical hope is that logic should be able to act as a neutral arbiter of metaphysical disputes, at least as a framework on which all parties can agree for eliciting the consequences of the rival metaphysical theories. An obvious problem for this hope is the proliferation of alternative logics, many of them motivated by metaphysical considerations. For example, rejection of the law of excluded middle has been based on

2 Pelletier et al. (2008) call anti-Platonist realists '*physicalists*'. We opted for 'structuralist', because it is not clear that the contemporary defenders of that position believe that these most general structures are physical. In any case, the main contrast is that Platonists believe that logic describes objective facts in a "third realm", while structuralists/physicalists believe these facts to be in the same world as the physical facts. We will come back to this issue in Chapter 10 when discussing anti-exceptionalism about logic.

3 Gottlob Frege should be listed here too. Frege was certainly a realist and he is well-known for his platonism and his view that logic and mathematics do not belong to the physical realm, but to a "third realm".

> metaphysical conceptions of the future or of infinity. Quantum mechanics has been interpreted as showing the invalidity of one of the distributive laws. Dialetheists believe that paradoxes about sets or change manifest black holes of contradiction in reality itself. There is no core of universally accepted logical principles. (Williamson, 2013, 212)

But perhaps that is not a bug but a feature of logic. Instead of trying to avoid metaphysical commitments, one might just as well endorse them and defend the specific metaphysical content of logical theory. The law of excluded middle, for example, is then true because it correctly represents the metaphysical facts.

A second possible way to arrive at such a structuralist view could depart from Ted Sider's assumptions about the joint-carvingness of logical connectives, which we encountered in Chapter 4. On Sider's view, logical constants are themselves *representational*. Their meaning may or may not "carve at the joints" (Sider, 2011).

As we also learned in Chapter 4 (and discuss again in Chapter 10), Sider does not believe that *metalogical* concepts, such as the concept of a logical constant or that of logical truth or consequence, carve at the joints. But Sider's reason for denying that metalogical concepts carve at the joints is their alleged impotence to "improve our fundamental understanding of the world" (Sider, 2011, 223). It's not clear that this view sits comfortably with the idea that some logical constants carve at the joints. After all, if some logical constants carve at the joints, then some sentences in joint-carving terms will exemplify the logical form of necessary truths (and thus constitute the class of what we typically consider to be logical truths, such that all their substitution instances will also be necessary truths). If the structure of the world is indeed such that it gives rise to this phenomenon, then having a theory about it would surely improve our fundamental understanding of the world.

However, if we assume the metalogical vocabulary to be also joint-carving, then logical theory would again be a very general metaphysical theory, another instance of structuralist realism about logic.

Arguments for Realism

In characterizing examples for realist positions, we already encountered reasons for holding the view. If you believe that the subject matter of logic comprises abstract objects (e.g. propositions and relations between them) *and* believe that they exist and have their logical properties independently of us, then you are committed to logical realism. Also, if you believe that logical theory (inevitably) makes certain very general (metaphysical) claims, *and* that these are true or false independently of us, you are a realist. In both cases, realism is a consequence of a prior commitment to certain kinds of objective facts that one takes to be the truth-makers for the claims of logic.

But are there also arguments in the other direction? Arguments that establish that there are objective, mind-independent logical facts in the first place? As we alluded to above already, the fact that we are inclined to say that claims about validity are true or false doesn't yet establish realism (at best it establishes only one of two merely necessary conditions). Neither does the fact that realism seems to be the "default position". As Michael Resnik (2000) points out, it's not even clear in what sense and where logical realism actually is the default position. In comparison, many working mathematicians and scientists are *mathematical* realists. This seems to give some weight to mathematical realism, and at least somewhat shifts the burden of proof to the mathematical anti-realist. But the same can't be said about logical realism. It's not clear that many scientists are logical realists in that sense – at best, logical realism is the default in *philosophy of logic* (Resnik, 2000, 184).

More promising arguments for logical realism have been provided by Stewart Shapiro (2000). Shapiro invokes criteria developed by Crispin Wright (1992) in order to determine whether the discourse about logic should receive a realist interpretation. One such criterion is whether the discourse in question is "epistemically constrained". A discourse is *epistemically constrained* if the possibility of unknowable truths in the discourse can be ruled out *a priori*. Wright argues that if a discourse is *not* epistemically constrained, then there is also no anti-realist interpretation of that discourse available. The idea is that if it doesn't depend on us what is true or false in a discourse, it must be tracking something real and objective.

Thus, if it could be shown that our discourse about logic is not epistemically constrained, realism would have a strong case. But now we need to be careful what we take to be the relevant discourse. First of all we should distinguish between the object-language logical truths themselves and metalanguage statements about them. In other words, we should determine, whether the anti-realist is denying the objective truth of

(5.1) Snow is white or snow is not white.

or denying the objective truth of

(5.2) 'Snow is white or snow is not white' is logically true.

Let us call an anti-realist who only denies the objective truth of statements like (5.2) a *status anti-realist*. Michael Resnik (2000) defends such a position. He considers truths like (5.1) to be objective, but argues against the objective truth of claims like (5.2).

We should further distinguish denials of the objectivity of (5.1) and (5.2) from denials of the objectivity of certain statements of formal (meta-)logic as (5.3) and (5.4).

(5.3) $p \vee \neg p$

(5.4) $\ulcorner \phi \vee \neg\phi \urcorner$ is a tautology in propositional logic.

Shapiro argues that anti-realism about statements like (5.3) or (5.4) is not plausible. If you take second-order logic to be logic, then there are truths of second-order logic (thus, statements like (5.3)) that are not provable and hence not knowable (because of incompleteness). Shapiro also argues that anti-realism about statements like (5.4) should already be implausible for first-order logic, since first-order consistency is not decidable, and hence there are metalanguage truths about the consistency of first-order sentences that are not knowable (as a result of a mechanical procedure).

Now, you might consider second-order logic not to be logic (for the reasons discussed in Chapter 3) and not be worried by undecidability – after all, the fact that it isn't *provable* that a certain sentence is consistent doesn't mean that it's not knowable that it is. You might have good reasons to believe it's consistent that fall short of a formal proof.

But, more importantly, you might restrict your logical anti-realism to informal logic. Claims like (5.4) belong to mathematics – or so you might argue – and thus may be fully objective.[4]

[4] In fact, that is the view endorsed by Resnik (1999).

Crispin Wright offers three more specific criteria that Shapiro applies to the case of logic: "width of cosmological role", the "Eutyphro contrast", and "Cognitive Command". Unfortunately, the application of these criteria is somewhat indecisive.

Width of cosmological role refers to the idea that a discourse is apt for a realist construal if statements of that discourse also feature in explanations outside that discourse. Now, logical implication and logical inference play a role in explanations almost everyhwere. Famously, the *deductive-nomological* model of explanation declares it a central feature of explanations that the *explanandum* logically follows from the *explanans.*[5] But that is not quite what is required for width of cosmological role. It seems that what we need is that statements *about* logical inferences or logical truth (statements like (5.2)) play a role in explanations outside logic. But do they?

Perhaps when explaining why we find certain arguments particularly compelling (a fact about human psychology, say), it might be explanatory to learn that these arguments are logically valid. Although this would be explanatory, the fact that the discourse of the *explanandum* is now that of human psychology of course doesn't speak in favour of logic's independence of human psychology (neither does the fact that statements about logical implications play a role in explanations in, say, semantics).

Wright's second and third criteria can be discussed together. The *Eutyphro contrast* is the contrast between reading biconditionals like (5.5) in a *detecting* or *constituting* way:

(5.5) An argument is valid iff an ideal agent acting under optimal conditions judges it to be valid.

Does the judgement of the ideal agent *make* the argument valid? Or does the agent merely *detect* that it is valid? This contrast is Wright's second criterion. If it's the ideal agent's judgement that makes the argument valid, this would speak for anti-realism.

[5] The *deductive-nomological model of explanation* (DN-model, also known as the Hempel–Oppenheim model, after Carl Gustav Hempel and Paul Oppenheim, who developed the model) specifies that scientific explanations of single (non-chance) events proceed by subsuming the event under a "covering" law. This is achieved by deducing a statement describing the event to be explained (the *explanandum*) from a statement of the relevant law of nature and a statement of background conditions. The law and the background conditions together form the *explanans*.

The criterion of *Cognitive Command* is satisfied by a discourse iff it is *a priori* that disagreements in that discourse must be attributed to a cognitive shortcoming of at least one of the disagreeing parties. In other words, if you disagree with your friend on a subject matter that satisfies Cognitive Command, at least one of you must have made a cognitive mistake.

But, as Shapiro concedes, both the Eutyphro contrast and Cognitive Command are not directly applicable, since both explicitly involve logical notions. For example, the "cognitive shortcomings" of Cognitive Command explicitly list inferential error. However, Shapiro suggests that the fact that valid arguments preserve truth might allow us to conclude nevertheless that (5.5) must be read in a detecting way (and, presumably, someone who misjudges an argument to be truth-preserving when it isn't must have made an objective error).

> The fact that logical consequence preserves objective truth allows for some objectivity. Suppose that we are dealing with an area of discourse [...] that is agreed to be objective. If an ideal subject under optimal conditions determines that an inference is valid, then she has detected that the inference preserves objective truth. No fiat and no non-cognitive stance can *make* an inference truth-preserving. (Shapiro, 2000, 362–363)

But that doesn't seem right. As Jody Azzouni (2014) explains, truth-preservation is cheap, once it is recognized that (i) the truth idiom is governed by the Tarskian biconditional

(5.6) $\ulcorner S \urcorner$ is true $\leftrightarrow S$

and (ii) one can't supplement either side of the Tarski biconditional with conditions that aren't equivalent to that side of the conditional. However, if that's observed, then for every set R of logical principles, one can supplement R with the following inference schema:

(**T**) $S \vdash T\ulcorner S \urcorner$, and $T\ulcorner S \urcorner \vdash S$

The introduction of this inference schema preserves consistency and "makes" R truth-preserving (regardless of R's consistency). If $U \vdash V$ according to R, then

(5.7) $U \vdash V \leftrightarrow T\ulcorner U \urcorner \vdash T\ulcorner V \urcorner$

holds in [R, $\mathbf{T}$] (Azzouni, 2014, 44).[6] Thus the arguments for logical realism seem inconclusive. Let us then look at the alternatives.

Psychologism

The term '*psychologism*' entered into the philosophical vocabulary as a pejorative label for philosophical positions which confuse or conflate non-psychological with psychological entities. The term was coined during the *Psychologismusstreit* (the "psychologism dispute") in German-speaking philosophy between 1890 and 1914 (Kusch, 2015). The most prominent opponents of psychologism were the (Platonist) realists Gottlob Frege and Edmund Husserl. On the side of the proponents, things are less clear. One of the main targets of the discussion was, in any case, John Stuart Mill and his *System of Logic* (Mill, 1843). However, it seems controversial among Mill scholars to what extent he actually held psychologistic positions (Kusch, 2015).

In the historical debate (the *Psychologismusstreit*) several questions were conflated. Is logic a sub-branch of psychology? Is logic used in psychology? Are the laws of psychology necessarily vague and imprecise? (See Kusch, 2015.) Some of these questions (especially those about the disciplinary boundaries of the science psychology) don't matter much for the purposes of this chapter. Thus, we suggest considering *Psychologism* to be the metaphysical view that logic is in an important way constituted by facts of (human) psychology or cognition.[7] The nature of this constitution can differ. Pelletier et al. (2008, 7–8) suggest distinguishing between four different types of psychologism about logic:

- *Psychological Individualism*: Identify logic with how individual people cognize about logic. Doing this on an individual-by-individual basis would

[6] Also note that, for any area of discourse that we regard as objective, there will be several distinct logics that preserve truth in that discourse.

[7] John Stuart Mill (1806–1873) Although it is controversial to what extent Mill was committed to psychologism, he made some claims that seem to identify the constituting facts of logic with facts about human psychology:

"I consider [Non-Contradiction] to be, like other axioms, one of our first and most familiar generalisations from experience. The original foundation of it I take to be, that Belief and Disbelief are two different mental states, excluding one another. This we know by the simplest observation of our own minds." (Mill, 1843, II.vii.5) possibly yield differing accounts of logic depending on the individual being examined.

- *Psychological Descriptivism*: Identify logic with some description of the observable performance of people's behaviour in the realm. For example we might consider how all the individuals cognize about logic, or perhaps how some important subgroup cognizes about it, or perhaps what the publicly observable output of these people is. We include here various ways to bring in the social realm in which people write about or discuss logic and mathematics.
- *Cognitive Architectures*: Argue that whatever might be discovered about logic is a function of the human cognitive machinery that is doing the discovering, and that this machinery itself has been shaped by the world in which it must operate. Postulate an architecture that in some way reflects some important aspect of the logical and mathematical realm. Identify logic with whatever is given by some "faculty" or "module" within this architecture that is common across people.
- *The Ideal Cognizer*: Identify logic with what an ideal cognizer would cognize about logic. This is on a par with "ideal observers" in other philosophical theories such as ethics.[8]

Psychologism in all these forms can be motivated from anti-platonist and naturalist considerations. If there are no abstracta, then logical truths must describe something concrete, perhaps concrete inferences that take place in the minds of reasoners. According to typical naturalist assumptions, we can have knowledge of anything only if that subject matter is part of the natural world. Again, the inferences of humans, and their cognitive architecture is part of this world. As we have seen above, structural realism would also offer a conception of logic that would satisfy anti-platonists and naturalists in this regard. However, one might consider it more plausible that logic describes after all the laws of thought rather than some very general metaphysical laws (especially if one is motivated by anti-platonist or naturalist concerns). But, as we said already, psychologism came under heavy fire.

[8] The characterization of these types is quoted from Pelletier et al. (2008, 7–8) but here formulated as claims about logic specifically.

Objections to Psychological Individualism

Psychological Individualism is what most opponents of psychologism, and Frege and Husserl in particular, seem to have in mind. Most of the arguments that Frege, Husserl and their contemporaries levelled against psychologism are directed at this variant and don't apply to other versions of psychologism. Psychological Individualism holds that the meanings of logical terms are subjective ideas in the minds of individuals. Likewise, logical laws or principles of logic describe very general psychological laws that govern the reasoning operations of individual human minds. Pelletier, Elio and Hanson (2008) identify an impressive number of 18 objections in the literature against this form of psychologism. We will rehearse some of the most important here.

The first objection starts with observing that Psychological Individualism leads to a form of relativism. All claims about logic will be claims about one individual's internal psychology. But then there can be no objective logical truths, only subjective ones.

Also, if the laws of logic are describing contingent facts about one individual, then these laws are neither necessary, nor do they have any normative force. Note that both hold for the individual as well as in general. (Why should everyone reason in this way, if Barbara reasons this way? Why should even Barbara reason this way, if it is just a contingent descriptive fact about her that she does?) Of course, even if all reasoned in the same way, neither necessity nor normativity would follow from that either.

Finally, everyone would have their own logical objects. If we suppose the meaning of the logical constants to be individual ideas, then there would be Jim's negation and John's and Barbara's, etc. (Frege (1980) famously ridicules the view that numbers are just individual ideas by pointing out that then there would be as many number 2s as there are individuals with such an idea.) Perhaps more problematically, this seems to lead to a puzzle about communication: how can we share thoughts in communication, if the contents of our terms are individually different – how would it even be possible to agree or disagree about something? Let's suppose that the law of excluded middle has a subjective content for Jim and a subjective content for Barbara. If Jim and Barbara both, apparently, assent to "it", then Jim is in fact assenting to his law and Barbara to hers. They are simply talking past one another.

Not all of these objections are equally problematic. Perhaps relativism isn't so absurd that the mere fact that a view entails it should already count against that view. Likewise, it *seems* that the laws of logic are necessary, but perhaps we are simply mistaken and they are not. Also, some of the objections can be avoided by changing from Psychological Individualism to Psychological Descriptivism. Psychological Descriptivism still holds that the output of our logic-related behaviour (the inferences we make, the ones we judge valid/invalid, the way we use logical terminology, etc.) is based on an individual psychological competence. But the *common content* of the relevant mental states or episodes is provided by the publicly observable performance or its outputs (Pelletier et al., 2008). Communication then succeeds if the individual mental states or events are "similar enough", but don't need to be identical. That they become similar enough is perhaps negotiated in a community via the publicly observable outputs (the arguments we produce, the claims about validity/invalidity we make, etc.). Thus, identity of content is not necessary for successful communication, and the actually necessary overlap of content could emerge as a tacit convention. We will discuss this view below in the subsection on conventionalism.

Instead of binding the individual contents together via social processes (like those that lead to the emergence of tacit conventions), one might also speculate about the psychological implementation of our knowledge of logic. Perhaps our mind/brain is sufficiently constrained (as a matter of biology) such that the individual "logics" we can develop are limited in a way that guarantees sufficient common content. Such a view, which would fall into the Cognitive Architecture category, has recently been defended as a neo-Kantian account of the metaphysics of logic. We will discuss this view also in a subsection of its own below.

Finally, one could introduce a notion of an "ideal cognizer" and take logic to be determined or constituted by what such an ideal cognizer would cognize about logic. Perhaps this could deal with the normativity problem (it is because the ideal cognizer is cognizing under *ideal* circumstances that everybody else *should* cognize that way too).

Pelletier et al. (2008) see two difficulties with that notion. One is that there is a danger that the notion is circular. We want to ground logic in whatever it is that this ideal cognizer cognizes, so we can't define what we mean by 'ideal' with the requirement that the ideal cognizer gets logic right nor should we define the circumstances in a "whatever it takes"-way.

So ‘ideal’ will have to refer to specific circumstances under which the ideal cognizer is cognizing. A second worry is that the ideal cognizer is itself an abstract concept, but the (or at least one important) motivation for psychologism is to avoid placing the facts that constitute logic in a “third realm” – but now our ideal cognizer seems to be dwelling in that very realm.

Perhaps that last worry can be met. We could understand the ideal cognizer to be an idealization of actual reasoners. The ideal cognizer represents how *we* would cognize about logic under ideal circumstances (and this might be grounded in concrete facts about our logical competence and our dispositions to employ it under specific ideal conditions). Again, this could help with the intrapersonal normativity problem (I should reason in this way, because that’s what my ideal self would do), but not with the interpersonal one, or the problem of relativism. But presumably a combination of the last two approaches could do that job: an ideal cognizer who is constrained by the architecture of his human mind/brain. Such a cognizer would be a role-model for all of us (humans, in any case).

Neo-Kantianism

Kantian constructivism in logical theory, or KCLT for short, is the view that “logic is the result of the constructive operations of an innate cognitive capacity that is necessarily shared by all rational human animals, and governed by categorically normative principles” (Hanna, 2006b, 68). This view has recently been defended by Robert Hanna, as an alternative to both psychologism and realism. On Hanna’s view there is a mutual constitutive relation between human rationality on the one hand and logic on the other.

Human beings are a special bunch. Sure, they are animals, i.e. sentient living organisms, but what sets them apart from all other animals is their *rationality*. Animals that are rational in Hanna’s sense are:

- rule-following
- intentional (possess capacities for object-directed cognition and purporsive action)
- volitional (possess a capacity for willing)
- self-evaluating
- self-justifying
- self-legislating
- reasons-giving

- reasons-sensitive
- reflectively self-conscious
- in possession of concepts expressing strict (logical, epistemic, and deontic) modality.

According to KCLT, all and only animals that are rational in this sense cognitively construct logic. However, the range of logics that can be constructed by those rational animals is constrained. There is only a certain class of logics that humans can make sense of. What these logics have in common is – what Hanna calls – a *protologic*. This protologic seems, at the same time, to be the common core of all deviant logics that satisfy certain conditions for logicality[9] *and* serves as the metalogic in our study and evaluation of those deviant logics as well as in our translation and interpretation practices. The protologic is a constraint on our reasoning, which is, however, consciously accessible to us. That's why the protologic is unrevisable (it is always presupposed in our thinking about logic) and also *a priori*.

Hanna suggests thinking of the protologic as a logical component of Chomsky's generative universal grammar (UG). The idea is that humans possess an innate "language faculty" which they share with all other humans. This language faculty is able to generate grammars, those of the languages that can be acquired in a normal way during the time-span of a child's language acquisition. Hanna interprets Chomsky as holding that the protologic is part of UG. (See Collins (2009) for discussion.)

Hanna also wants to defend the view that logic is intrinsically normative and that it should be seen as a system of categorical imperatives. This system provides the formats or structures for the maxims by which we ought to reason. It seems that one of Hanna's motivations for considering

[9] *Hanna's conditions for logicality*:

(I) *The formal logic condition*. The non-classical logic (NCL) is a science of the necessary relation of consequence.

(II) *The representational adequacy condition*. The NCL's proposed extension of, or deviation from, classical logic is based on its being able accurately to represent in the format of symbolic logic some apparent linguistic facts that are not represented within classical logic: for example, strict implication or modality, constructibility of proofs, relevance, vagueness, future contingency, nonexistent objects, paradoxes, and so on.

(III) *The localization of application condition*.The NCL's scope of application is restricted to all and only those language domains containing the apparent nonclassical linguistic facts that it represents. (Hanna, 2006b, 77) logic to be such a system of canonical imperatives (rather than a system of hypothetical imperatives) is to avoid Quine's criticism of conventionalism. As we shall see below, Quine argued that a set of conventions can't generate logic, since you need a logic in order to generate anything from the conventions. However, if the imperatives of logic are categorical, then they are in a certain sense "inescapable". This would be a reason to favour neo-Kantianism over conventionalism, if Quine's criticism of the latter was convincing – we will discuss this in the next section. (The nature of logic's normativity will be discussed in detail in Chapter 8.)

A potential problem for this view is the assumed status of protologic. As Hanna explains, protologic is supposed to be structurally different from all object logics that satisfy the criteria of logicality but can also serve as a metalogic (since it is the logic from which we compare and evaluate all other logics). But it is not clear that the system we consider metalogic has the properties that protologic is supposed to have. Since protologic is considered to be presupposed by all logics that satisfy the requirements for logicality, it should be a neutral arbiter between logics. However, as Timothy Williamson (2013) has argued, metalogic is not neutral in that sense. From a structural point of view, metalogic is on a par with the object logics, but then either there isn't only one protologic, or, if there is only one that is perfectly intelligible to the human mind/brain (because it is hard-wired into the language faculty), there is also only one true logic (or protologic isn't the metalogic).

Conventionalism

For the logical positivists, Kant's conception of the synthetic *a priori* was in need of revision. There are no synthetic *a priori* truths, but instead conventions that are accepted in the adoption of a framework on the basis of pragmatic considerations. The recent history of science had made it clear to the positivists that "possible experience" is less constrained than Kant had anticipated. What seemed to be *a priori* for previous generations was rejected by contemporary scientists. Famously, Rudolf Carnap endorsed such a view also for the truths of logic. We freely choose a logic as part of a linguistic framework. In so choosing we can't get it wrong (or right, for that matter), since the choice is purely pragmatic. Questions about the true logic do not arise, since these would be framework-external and meaningless. The necessity of logic, on the other hand, is explained by the analytic status of statements about logic within the framework. Logic – all of it – is true by convention.[10]

There are several observations that support conventionalism and make it attractive. However, one *bad* argument is that logical facts (about what implies what, and which sentences express logical truths) seem to a large extent to be a matter of the meaning of certain expressions (the logical connectives). What meaning our words have is a matter of convention – or so one might argue – hence logic is a matter of convention.

This argument seems confused. True, it is a matter of convention that the English word 'and' expresses conjunction. If conventions were different such that 'and' expressed disjunction, 'Snow is white and snow is not white' would express a logical truth instead of a logical falsehood. But that doesn't show that logic is true by convention. It only shows that we could have chosen different words to express conjunction or disjunction.

But there are better arguments for conventionalism. We already observed in Chapter 4 that one might consider it a matter of convention which expressions of the language we treat as logical constants (those we "hold fixed"). Thus, at least the extension of 'logical truth' and 'logical consequence' is conventional depending on what choice we make there. Also, we have learned in Chapter 2 that there are several alternatives to classical logic. Many are well studied, even with fully developed applications to mathematics or science. It is easy to imagine that we could have ended up using such an alternative system as a logic (assuming for the moment that we are in fact using standard logic). It is also easy to imagine that at some point in the future we might choose to change our logic to one of those alternatives. It is very plausible to think that that would just constitute a change in conventions (Azzouni, 2014, 32).[11]

10 *Carnap's Principle of Tolerance:*
"It is not our business to set up prohibitions, but to arrive at conventions …
In logic, there are no morals. Everyone is at liberty to build up his own logic, i.e., his own form of language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments" (Carnap, 1937, 51–52).

11 You might wonder why *this* argument is better than the one we just rejected as "confused". Why are these other logics not in the same boat as a different version of English in which 'and' expresses disjunction? We will discuss this question in Chapter 7.

It is widely believed that Quine showed conventionalism to be flawed. Quine's most famous argument appears in his paper "Truth by convention" (Quine, 1976). In this paper Quine discusses whether logic and mathematics could be founded on conventions alone. Couldn't we just stipulate logic into existence by laying down a set of conventions? Remember how simple it is to lay down the proof theory of propositional logic. Just take the example of an axiomatic approach: we simply stipulate three axiom schemata and a rule of inference – enough to generate all truths of propositional logic. Doesn't that show that logic can be legislated true by convention?

In order to see what the problem is we will use a slightly simpler example from Lewis Carroll's beautiful paper "What the tortoise said to Achilles" (Carroll, 1895).[12] Let us assume we are interested in the following inference:

(A) If the two sides of this triangle are things that are equal to the same then they are equal to each other.

(B) The two sides of this triangle are things that are equal to the same.

(Z) The two sides of this triangle are equal to each other.

How can we formulate conventions such that they would take us from premises (A) and (B) to (Z)? It seems we need something like

(C) For all *x*, *y* and *z*, if *x* and *z* are true statements and *z* is the result of putting *x* for '*p*' and *y* for for '*q*' in 'If *p* then *q*' then *y* is to be true.

In other words, if ϕ is true and $\ulcorner$if ϕ then $\psi\urcorner$ is true, then ψ must be true. How will that help us in deriving

(D) (Z) is to be true.

from (A) and (B)?

Let us first simplify (A) and (B) into

(E) (A) and (B) are true and (A) is the result of putting (B) for '*p*' and (Z) for '*q*' in 'If *p* then *q*'.

With a further convention that would allow us to drop universal quantifiers and replace variables with constants, we could get from (C) to

[12] In fact, we will use a mix of Carroll's and Quine's examples.

(F) If (A) and (B) are true and (A) is the result of putting (B) for '*p*' and (Z) for '*q*' in 'If *p* then *q*' then (Z) is to be true.

Now we need to infer (D) from (E) and (F). But this is an inference for which (C) is needed again. From

(G) (E) and (F) are true and (F) is the result of putting (E) for '*p*' and (D) for '*q*' in 'If *p* then *q*'.

we are supposed to derive that (D) is true. But now inferring that (D) must be true from (G) and (C) is exactly analogous to our initial task of deriving (D) from (E) and (F) – we entered into an infinite regress.

The problem is that we need logic in order to apply the conventions. Without logic, the logical constants are supposed to mean nothing to us. Only after the conventions are in place is 'if *p* then *q*' supposed to acquire a meaning that would allow us to use *modus ponens*. But without *modus ponens* we can't apply the convention to our inference from (A) and (B) to (Z).

This is often taken to prove that logic can't be just true by convention (cf. Miller, 2007; Sider, 2011). But, as Jody Azzouni (2014) points out, Quine actually argues for a *dilemma*. We think of the conventions that constitute the foundation of logic either as *explicit* conventions or as *tacit* conventions. If we consider them as explicit conventions then we run into the difficulty just explained: we would need to already have a logic in order to know how to apply those explicit conventions.

But we can think of the conventions that ground logic as being *tacit*: conventions that are adopted through behaviour and that are formulated explicitly only later. Quine seems to have two objections against this second horn of the dilemma. The first objection is that it is not clear how we are supposed to determine what the conventions actually are, if we have to read them off the behaviour. When conventions are explicit, we can compare the behaviour with the explicit convention and determine when a behaviour is following the convention and when it is violating it. But how are we supposed to make that difference if all we have is the behaviour? Quine's second objection is why we should call tacit conventions "conventions" rather than firmly held beliefs (since it seems that – as far as behaviour goes – they will be identical in their observable outputs).

As Azzouni (2014) argues, both of Quine's challenges can be met. On the one hand, we have learned much more about how to conceptualize tacit conventions since Quine. Thanks to the work of David Lewis (1969), Ruth Millikan (1998) and Tyler Burge (2007), we know much better how to think of tacit conventions and how to distinguish them from other forms of systematic behaviour.

David Lewis (1969) shows that conventions can emerge without explicit agreement. Lewis defines a regularity *R* in action or belief as a convention in a population *P* iff, within *P*, the following six conditions[13] hold:

(C1) Almost everyone conforms to *R*.

(C2) Almost everyone believes that the others conform to *R*.

(C3) This belief that the others conform to *R* gives almost everyone a good and decisive reason to conform to *R* himself.

(C4) There is a general preference for general conformity to *R* rather than slightly-less-than-general conformity – in particular, rather than conformity by all but any one.

(C5) There is at least one alternative *R*' to *R* such that the belief that the others conformed to *R*' would give almost everyone a good and decisive practical or epistemic reason to conform to *R*' likewise; such that there is a general preference for general conformity to *R*' rather than slightly-less-than-general conformity to *R*'; and such that there is normally no way of conforming to *R* and *R*' both.

(C6) (C1)–(C5) are matters of common knowledge.

Burge (2007) argues that the requirements on common knowledge in this definition are even too strong. Participants in a convention don't need to know that they are participating in a convention, or that there are alternatives to their practice. What remains as one of the crucial factors that distinguishes behaviour based on tacit convention from other forms of robust behaviour is that it is an imitation of behavioural patterns that proliferate "due partly to weight of precedent, rather than due [. . .] to their intrinsically superior capacity to perform certain functions" (Millikan, 1998, 3). Thus, for a behavioural regularity to be conventional, there must be good enough alternatives. Also, the organism adopting the pattern must

[13] (C1)–(C6) are quoted from Burge (2007, 32–33).

be able to adopt it via imitation, and be in a position to adopt the possible alternatives in the same way.

Against Quine's second challenge that we couldn't empirically distinguish tacit conventions from firmly held beliefs, we can respond that there is indeed empirical evidence that logic can satisfy the empirical criteria for conventionality. For example, we have evidence that how humans actually reason is not like logic – because we reason with topic-specific tools, rather than with an "all-purpose topic-neutral piece of algorithmic machinery" (Azzouni, 2014, 43). But that makes it plausible that logic emerged as a reasoning tool that we normatively impose on our reasoning practices.

As this discussion shows, conventionalism is still a plausible contender for a metaphysics of logic. At least, the typical argument against it, which is only based on Quine's regress, is insufficient to establish that conventionalism must be mistaken.

Non-Cognitivism

The anti-realist positions we have encountered thus far all rejected the second of Resnik's necessary conditions for logical realism, the *objectivity condition*. In the last section of this chapter we want to look at a position that rejects the first condition, the *cognitivism condition*.

Crispin Wright (1986) and Michael Resnik (1996) both argue for non-cognitivism about logic. Wright's motivation for non-cognitivism develops from considerations very similar to the *Cognitive Command* criterion we discussed above in the section on realism. Wright argues that it is always possible to have a genuine disagreement about the logical necessity of a given statement, without the disagreement being due to ignorance or error on either side.

Wright considers two subjects, X and Y, who are both presented with the following proof:

(1)	$A \to B \Rightarrow A \to B$	
(2)	$A \Rightarrow A$	
(3)	$A \to B, A \Rightarrow B$	1,2 MPP
(4)	$A \to B, A \Rightarrow B \vee C$	3, vel-I

Imagine that X follows the proof and accepts it as such, thereby accepting the logical necessity of the following conditional describing the proof:

> If any proof commences with a pair of assumption-sequents, $A \Rightarrow A$, and $A \to B \Rightarrow A \to B$, followed by the *modus ponens* step which those two lines furnish, followed in turn by a step of vel-Introduction on the result with C as the right-hand constituent in the then resulting disjunction, then that disjunction will be $B \vee C$, and will depend on A and $A \to B$ as assumptions. (Wright, 1986, 204)

Y, on the other hand, only regards the structure as a corroboration of the conditional, which he considers to be enormously probable but not logically necessary.

If matters of logical necessity (and related notions) were cognitive, then we should be able to point to some specific error or ignorance on the side of Y – what could that be? To suggest that Y doesn't understand the notion of necessity (and thus is making a conceptual mistake) might be unjustified if Y can prove to us that he does understand the concept. On the other hand, to suppose that Y simply lacks the ability to recognize a proof, would be a desperate move (since it could be made on the same grounds also in contexts that we *do* believe are non-cognitive, such as disputes over what's funny). However, if a reaction like Y's always seems possible, then logical necessity is not a cognitive notion.

Resnik's motivation for non-cognitivism is the fact that he finds all other anti-realist alternatives less plausible. According to Resnik (1996), judgements about logicality (logical necessity or validity) aren't true or false, and their function is not to state any facts. Instead they serve two other functions.

The first function is to signal how we want certain statements (in the scope of 'it is logically necessary that…', for example) to be treated in our everyday inferential practice. In this particular case, we might want to signal that our claim is not in need of further justification, or that everyone (rational and intelligent) should agree to it.

The second function is to describe commitment:

> In saying that A is logically true I commit myself to its truth; in saying that A implies B, I commit myself to its being true that if A then B; in saying some statements imply A, to at least one of them being false if A is; in saying that A is consistent with some statements, to denying that they are

> all jointly true only if they are also false. In each case, I commit myself at least in the sense that I can be criticized for uttering the logicality judgement engendering the commitment, in case the statement to which I am committed turns out to be false. (Resnik, 1996, 511)

To assign such functions to judgements about logicality might work (to some degree) for free-standing judgements. However, projectivist theories, like Resnik's non-cognitivism, have notorious difficulties in dealing with mixed contexts, i.e. contexts in which (alleged) non-cognitive statements occur embedded in cognitive statements – a difficulty also known as the *Frege–Geach Problem.*[14] Resnik is aware of the problem and discusses some examples, like the following:

(5.8) The conjunction of logical truths is logically true itself.

(5.9) Upon receiving Russell's letter Frege realized that his system was inconsistent.

Resnik agrees that these statements have truth-values, and thus suggests that the embedded use of logicality statements does become fact-stating. For example, he renders (5.8) as

[14] The Frege–Geach Problem is a problem that arises for all non-cognitivist accounts that want to deny that free-standing statements of the alleged non-cognitive discourse have normal truth-values.

To see the problem, consider a view that would deny that moral statements, such as 'Tormenting the cat is bad' are true or false. Such statements will, nevertheless, occur in arguments like the following (taken from Geach, 1965):

1. If tormenting the cat is bad, getting your little brother to do it is bad.
2. Tormenting the cat is bad.
3. Ergo, getting your little brother to torment the cat is bad.

Here the problematic statement occurs in the second premise, and is also the antecedent of the first premise. We would normally consider the argument to be valid, but if non-cognitivism about moral judgements is right, then the second premise doesn't have a truth-value and doesn't state a fact. But the first premise does state a fact and, presumably, has a truth-value. Does that mean that 'Tormenting the cat is bad' has a different meaning in (1) and (2)? If so, is the argument, after all, not valid, but merely an equivocation? Non-cognitivists are challenged to provide an account that solves this problem. See Schroeder (2008) for strategies to deal with this and related problems for a semantic theory along non-cognitivist lines.

(5.10) If we treat the conjuncts as logically true then we (should) also treat their conjunction as logically true.

On this reading, (5.8) is stating a fact about the norms that, supposedly, govern our inferential practices. But the challenge of the Frege–Geach Problem is not just to come up with some paraphrase or other for embedded use of the problematic vocabulary, but to come up with a systematic account that shows how this embedded usage fits into a compositional theory of meaning. This still seems to be an open problem for a non-cognitivist account of logic.

Questions

1. Provide examples for discourses that are epistemically constrained, i.e. discourses for which the possibility of unknowable truths in the discourse can be ruled out *a priori*.
2. What is wrong with the following argument?
   (P1) Almost no discourse is such that we (already) know every truth in it.

   (P2) We can't rule out *a priori* that we will all die (simultaneously) in the next moment (say, because of a global catastrophe).

   (P3) If we are dead, then truths that we did not know when alive will remain unknown by us.

   (C) Therefore, there is almost no discourse for which the possibility of unknowable truths in the discourse can be ruled out *a priori.*
3. Consider again Quine's argument against conventionalism (the one that showed by example that you need logic in order to apply conventions). Does the example that was used in the argument show that we need to have a specific logic? Does it show that we need a full-blown logic (e.g. all of standard first-predicate logic)?

6 The Epistemology of Logic

In the previous chapter we discussed several different answers to the question of what it is that makes the laws of logic true (if they have truth-values at all). In this chapter we will look at the question of whether we know logical truths at all and – if we do – how we get to have knowledge of them. Of course, the epistemology of logic is not completely independent of its metaphysics. Thus, some of the positions that we discussed in the last chapter will also feature in this chapter.

In *metaphilosophy*, the philosophy of philosophy, one often finds arguments that try to derive conclusions about the metaphysics (the subject matter) of a certain area of philosophical enquiry from observations about the methodology of that enquiry. The idea behind such an inference seems to be this: a certain methodology would not make sense if the subject matter wasn't such and such, therefore it is the best explanation of the observed methodology in this discipline that it is aimed at such and such a subject matter (see Goldman (2007) for an instance of this inference). So it seems that if one just carefully studied the methodology (and thus the epistemology) of logic, one could find answers to the questions discussed in the previous chapter, the metaphysics of logic.

But there are two things to note that should diminish the hope one might have in such an inference from epistemology to metaphysics. The first observation is that there is no guarantee that philosophers are using the best or even an appropriate methodology for studying the subject matter of their discipline. Since the subject matter is contested (after all, we are looking for an argument that is supposed to enlighten us on what the subject matter actually is), philosophers might have picked the wrong methodology because they had false beliefs about the nature of the subject matter. But philosophers might, of course, also have picked the wrong methodology despite not having false beliefs about the subject matter.

Perhaps they made a collective mistake or adopted a methodology out of a tradition without sufficient critical reflection (which is what many critics of philosophy's armchair methodology believe to be the case). To hold that the methodology is appropriate (even though beliefs about subject matter might not be) would be justified if one could point to progress in philosophy that would seem to be miraculous if the methodology was inappropriate. Unfortunately, for most areas of philosophical enquiry, the progress is so modest that this argument seems weak (see Chalmers, 2015).

One might counter with the observation that logic, at least, is doing much better than most other philosophical disciplines here. Logic is making progress, hence, whatever it is that logic is studying, the methodology seems appropriate to that subject matter. This is where our second observation comes in. Actually, the *prima facie* methodology of logic does not discriminate clearly between different accounts of the subject matter – it is compatible with several of them. Let us quickly see why.

The *Prima Facie* Epistemology of Logic

By *prima facie epistemology* we mean a descriptive account of how we come to believe and how we justify claims about logical truth, logical implication, etc. According to Nelson Goodman (1955), we learn about the validity of particular arguments from subsuming them under general principles. On the other hand, we recognize the validity of these general principles, by recognizing that they give the right verdict about the validity of particular arguments. What holds for logical validity (or logical consequence) holds also for logical truth.

In case an argument that we thought was invalid is judged valid by the general principles we hold to be true, we have to make a choice about where to make an adjustment. We can either change our view about the validity of the particular argument, or amend our general principle. These adjustments are necessary until our beliefs about general principles and our beliefs about which arguments are valid are in "reflective equilibrium". In the process of making these adjustments we are guided by certain general principles of theory choice (e.g. simplicity, symmetry, entrenchment) as well as other beliefs we might have. If this epistemology about logic is only a part of a wider process of mutual adjustments between beliefs, we will talk of a "wide" equilibrium (how "wide" this process can be will be discussed below).[1]

Let us grant to Goodman that this is a (sketchy but) largely adequate picture of how we come to believe and how we justify claims about logical truth, logical implication, etc. It's worth emphasizing that this story is compatible with all the options we have discussed in the previous chapter. Let us first consider non-cognitivism. Nelson Goodman himself was a non-cognitivist about logic (see Goodman, 1955; Cohnitz and Rossberg, 2006). Michael Resnik (1999) is another non-cognitivist who holds that (a wide version of) reflective equilibrium best describes our epistemology in the case of logic. *That* non-cognitivism is compatible with this *prima facie* epistemology is hardly surprising. Non-cognitivism is compatible with almost any such epistemology; since there are no facts one is out to detect, no "epistemology" is going to be inadequate (at best, a very elaborate apparent epistemology could be a little bit too much ado about nothing).

With respect to psychologism, Kantianism, conventionalism or realism, the *prima facie* epistemology is just too general to discriminate between them. The way we described it, the "method" of reflective equilibrium has two main components. One is the mutual adjustment between believed general principles and judgements about particular instances. The second component is that the initial inputs into the process of mutual adjustments are antecedently held beliefs (about instances or principles).

One might perhaps think that the first of the components is better suited to a non-cognitivist account because of the involved circularity: general principles are supported by individual instances, individual instances are supported by general principles; if support is lacking there is a pragmatic choice about where to make the adjustment. But the same is also

[1] Although Nelson Goodman first described the process in *Fact, Fiction, Forecast* as the process by which we justify principles of deduction and validity judgements about inferences, the term 'reflective equilibrium' was coined by John Rawls in his *A Theory of Justice* (Rawls, 1971). Rawls suggests a reflective equilibrium between considered judgements and principles of justice.

It should be noted that Goodman offers the reflective equilibrium story as a descriptive account of how we in fact justify judgements about validity. Most authors who critically engage with Goodman's proposal mistakenly believe that he was offering a normative proposal for how we *should* justify beliefs about judgement-independent validity-facts. We will see below that Goodman had nothing of that kind in mind.

true for areas of enquiry where one might hold a robust factualist view. Whenever our (believed to be true) theory is in conflict with recalcitrant data, we have to make a choice whether to adjust the theory or to discard the data. It is not in general the best strategy to adjust the theory to the data. This will make the theory very complicated and ignore the fact that there are observation errors. But then one has to choose between simplicity and empirical adequacy, a process of mutual adjustments between particular data points and the general theory (cf. Quine, 1953a).

So, the first component is just a consequence of a general epistemic predicament. If the choices we make when confronted with recalcitrant data are pragmatic, then different choices might lead to different equilibria and hence to different theories (not all of which can be true). In other words, theories are underdetermined by data. This might sound frustrating, but it is a frustration that the factualist is familiar with from all areas of enquiry.

The second component is a bit more interesting, but unfortunately too unspecific to clearly favour a particular metaphysical account. The antecedently held beliefs that serve as inputs to the reflective equilibrium process can be accommodated by all accounts, with only slightly different stories. Conventionalists, psychologists and neo-Kantians, for example, might locate facts about logic in linguistic conventions that we adopt or acquire together with our linguistic competence (cf. Philipse, 1989; Hanna, 2006a), or which are part of our reasoning competence. In this case, the antecedently held beliefs about particular instances and general principles are perhaps simply outputs of this linguistic or reasoning competence, comparable to the intuitive judgements about grammaticality that (Chomskyan) linguists consider to be the relevant data for inquiries into grammar. In this case, the data (the input into the reflective equilibrium process) is the output of the very competence/disposition/set of internalized conventions, etc. that such an account would consider to be the relevant subject matter of logic (see Cohnitz and Haukioja (2015) for a detailed story along these lines). Realists, on the other hand, might consider the antecedently held beliefs to be the outputs of a special faculty of "rational intuition" which provides (defeasible) evidence about metaphysical facts or the platonic "third realm" (Bealer, 2000), or – if more naturalistically inclined – will consider the antecedently held beliefs to be very general, deeply entrenched empirical beliefs (Williamson, 2007).

In other words, all metaphysical accounts will have to tell some story about why the antecedently held beliefs that we take as starting points for the reflective equilibrium process are legitimately playing that role, but all have a story to tell. Sure, some of these stories are better than others, but even for the bad stories there are philosophers who explicitly subscribe to these, so the inference from epistemic practice to the implicitly assumed metaphysics (let alone actual metaphysics) of logic is blocked.

It should also be noted that these stories will come apart (to some extent) when it comes to the question of how to improve the existing methodology (and how to "widen" the reflective equilibrium). Neo-Kantians and other friends of psychologism might turn to additional evidence from psychological studies of human reasoning processes and cognitive science (Hanna, 2006a), linguistic conventionalists will turn to psycholinguistic evidence or corpus studies, structural realists will turn to fundamental physics, etc. However, most accounts seem to have learned to live with the standard methodology and just provide different interpretations of it.

Three Dogmas

In this chapter we want to focus on three primary questions:

1. Is logic revisable?
2. Is logic *a priori* or *a posteriori*?
3. Can we justify logic?

As usual, we will have to disambiguate these questions in the course of the chapter. Let us observe first that the three questions really are distinct. The first question asks whether logic can be revised. This is independent of the second question of how we can know about logic. Perhaps, if logic is *a posteriori*, it can be revised *a posteriori*. And if it's *a priori*, it can be revised *a priori*. But, also, logic might be *a priori* but unrevisable.

The third question asks about whether logic is justified. If you think that *a priori* and *a posteriori* are ways of knowing *and* believe that justification is necessary for knowledge, then if there is any positive answer to the second question (e.g. that logic is *a priori*), we will have to answer the third question also in a positive way (because if we know it's knowable *a priori*, then we know that it is knowable, and then we know that it is justifiable). But this is not how you have to think about these categories. Some philosophers of science (e.g. followers of Karl Popper's school of thought[2]) would hold that scientific theories are *a posteriori* revisable, but never receive positive justification. If we take that observation together with the previous one, then logic could turn out to be either *a priori* or *a posteriori* revisable, but not justifiable either way.

Although these questions seem to be distinct, we will see that they are interrelated in various ways. The three questions have standard answers, or – at least – answers that were mainstream until the middle of the last century. This *standard account* goes as follows: logic enjoys a special epistemic status, the strongest form of certainty: logic is self-evident. Our certainty is based in the *a priori* (because that is how you recognize self-evidence), and because it is *a priori* knowledge, it also enjoys ultimate justification. This standard account was shaken up considerably by Willard Van Orman Quine in his famous paper "Two dogmas of empiricism" (Quine, 1953a) and by Nelson Goodman in *Fact, Fiction, Forecast* (Goodman, 1955).

Let us begin with Goodman's contribution. He attacked the last component of the standard account, the idea that logic is justifiable. In *Fact, Fiction, Forecast* Goodman was concerned (among other things) with the problem of induction.

The classical problem of induction is the problem of how inductive inferences are justified, while justification, as Goodman says, is supposed to be very different from just describing how we actually make inductive inferences. But why do we require such a justification in the first place? One possible line of thought could be, that the end-results of inductive inferences are beliefs that we consider candidates for knowledge. In the sciences this seems most obvious: scientific knowledge, as the intended end-product of scientific enquiry, for a good deal seems to rest on inductive methods. It seems to rest on inferring generalized, universal statements (such as 'All emeralds are green.') from observed particular instances (all observed

[2] Sir Karl Raimund Popper (1902–1994) argued that scientific theories can never be verified, because their empirical content transcends whatever finite empirical evidence we can pile up in favour of the theory. However, theories can be falsified. We can find empirical evidence that is in conflict with the theory and decide to consider the theory refuted. For some philosophers in that school of thought, this implies that we can never have positive reasons to believe a theory, but only reasons to discard one. Science makes progress through conjectures and refutations, but we are never justified in claiming that our current theories are true.

emeralds in the past have been green), and on predictions about future events from the observation of events in the past. Now, in order to make sure that our beliefs in generalized universal statements have any chance to qualify as knowledge, they at least need to qualify as justified on the basis of our evidence, which in turn means that the inference we made from our evidence to the generalized universal statement must have been justified. But how could such an inference be justified?[3]

In order to answer this last question, we first need to know what we require of a justification. What seems traditionally to be required from a justification is a good argument that establishes that using inductive inferences does not lead us astray. We do not just want to know what we do when we make inductive inferences, but to know why it is *right* to make such inferences.

Although somehow it seems to be a meaningful question whether there is such a justification for our inductive practices, David Hume showed that there can be no such argument, if we reflect on what it means to be a "good argument", and what it means "not to lead us astray". A "good argument" is at least a valid and non-circular argument. For Hume, there are inductively valid and deductively valid arguments. Since inductive validity is under

[3] Goodman discusses the "old" problem of induction and the problem of deduction in *Fact, Fiction, and Forecast* primarily as stage setting for his *new problem of induction*, which is perhaps the most famous part of the book. The new problem of induction is the problem of how to chose the predicates that we want to use for inductive inferences, since it depends on our choice of predicates which generalizations we will consider supported by the evidence we have collected. Let us assume that it is before some time t and we so far have observed only green emeralds. Should we take these observed emeralds to support the generalization 'All emeralds are green', or the generalization 'All emeralds are grue', where 'grue' is defined as follows:

x is grue iff x is observed before t and green or is not so observed and blue

Since all observed emeralds are also grue, the past observations would also support the second generalization.

One might think that 'green' should be preferred over 'grue' because 'green' is simpler and doesn't contain a reference to time in its definition. But this, as Goodman argues, is a purely language-relative fact. If we had started with a language with 'grue' and 'bleen' in it (where 'bleen' refers to objects that are blue and observed before t or are not so observed and are green), then 'grue' and 'bleen' would be "simpler" and 'green' and 'blue' defined with reference to time, as in

x is green iff x is observed before t and grue or is not so observed and bleen scrutiny here, all inductive arguments for the justifiedness of induction must be disqualified for being rule-circular: the argument uses an inference in order to establish the justifiedness of that very same inference. This leaves us with deductive arguments for the justifiedness of induction.

If 'using inductive inferences does not lead us astray' means that inductive inferences are necessarily truth-preserving, then a deductive argument cannot establish that. The reason is that the content of the conclusion of an inductive inference transcends the content of the premises, and that means in turn that there can be no deductive route from the premises to the conclusion of an inductive inference. Although adding a premise in order to make an inductive inference deductive (e.g. a premise that claims the uniformity of nature) could be found, it would, if true at all, again be only either logically true or synthetically true. If it were logically true, then the conclusion of an inductive inference would also have followed deductively from the premises without it. Since, as we have seen, it doesn't, the extra premise can't be a logical validity. If, however, it is a synthetic truth, then it must itself rest on inductive reasoning, and again violates the requirement of non-circularity.

It is important to understand that Hume's argument is general. It is not just an argument against a particular attempt to justify induction in the sense above, but a general argument that there can be no such justification at all. In order to see the generality of this argument, Goodman showed that the same problem also arises for deduction. Various authors have disputed that the problem is the same, and Goodman does not spend much time in the text on why he thinks it is, so we will deliver the missing argument here.

Such an argument can indeed be found in Susan Haack's writings (Haack, 1996, 183), where she analyses the parallel between the dilemma for the justification of induction and deduction as follows. Hume's dilemma has two horns: the first is that deductive justification of induction would be too strong; the second horn is that a mere inductive justification of induction would be circular. Haack's dilemma for deduction has it the other way around: the first horn is that a deductive justification of deduction would be circular, while – and this is the second horn of Haack's dilemma – an inductive justification of deduction would be too weak.

Let us first look at the second horn of Hume's dilemma and the first horn of Haack's dilemma. That a deductive justification of deduction, or an inductive justification of induction, respectively, is problematically circular seems almost too obvious to require argument. (However, we will see below that there is some room for manoeuvre.) Typically when we accuse an argument or a line of reasoning of being "circular" we have a reasoning in mind that assumes as its starting point what is yet supposed to be established by the reasoning. In the most obvious case, such a reasoning would be in the form of an argument that has its conclusion also already as a premise. How exactly this form of vicious circularity or "question begging" is defined is a notoriously difficult problem. But this sort of circularity is not what we find in the arguments under consideration here. Instead of having the intended conclusion as a premise, the problematic arguments rather seem to rely on a rule of inference, the justifiedness of which is what the argument aims to establish. These horns thus object that an inductive justification of induction or a deductive justification of deduction would be *rule-circular*.

Let us now look at the other horns. That any purely deductive justification of induction is "too strong" is, of course, not to say that it actually proves more than it was supposed to prove (this could hardly constitute an objection). It is rather indicating a structural incompatibility between the means for establishing something and the claim that is supposed to be established. A purely deductive argument to justify induction is from the start bound to fail, since it is inadequate to establish anything short of logically necessary truth-preservation (and that inductive inferences are not truth-preserving with logical necessity is an analytic corollary of them being inductive). The reason is that a purely deductive justification can only get started from truths that are themselves logical necessities (e.g. axioms or theorems of deductive logic, or instantiations of these), and will by application of deductive inference rules always result only in logical necessities (deductive inference rules do not only preserve truth, they also preserve necessary truth). But no induction principle (and no claim about its justifiedness) is a truth of logic. Hence, no purely deductive justification could ever result in the right sort of conclusion.

One might wonder whether a "purely deductive" argument is indeed required, or whether there could be a deductive argument for a principle of induction that starts from non-necessary premises. However, in this case a problem similar to what we already encountered would arise: the premises of that argument would either transcend experience or not transcend experience. If they transcend experience, then they must already be based on a principle of induction (which leads us back to the circularity problem). However, if they don't transcend experience, then – since deductive inferences are always non-ampliative (i.e. their conclusion can't transcend the content of its premises) – the conclusion also won't transcend experience. But then no such argument could establish a general principle of induction.

In a very similar way any inductive justification of deduction is "too weak": showing that a certain inference usually holds is not the same as showing that it holds with logical necessity. Thus an inductive justification will not lead to a justification of deductive inference *qua* deductive inference. No inductive method will allow us to infer that a deductive inference rule is necessarily truth-preserving, i.e. that if the premises of such an inference are true, then necessarily the conclusion of that inference is true, too.

There is also a second problem with this horn. Let us assume for the sake of argument that in principle an inductive justification of deduction was potentially adequate. Then, as Paul Boghossian (2000) pointed out, any such inductive argument for logic would at some point *also* use deductive rules (at least in the metatheory). But if no inductive argument is purely inductive, then it will be just as circular as a purely deductive one. Hence, not only is induction unjustifiable; deduction is also unjustifiable. But if deduction is unjustifiable, then logic is unjustifiable too.

Quine attacks the first two components of the standard account. In "Two dogmas of empiricism" Quine is first concerned with the notion of *analyticity*. He reviews a number of traditional attempts to define what it means for a truth to be analytic and finds them all unsatisfactory – either the terms in which the *definiens* is provided are even less clear than the *definiendum*, or the definition is circular (crucial but problematic terms in the *definiens* are themselves defined in terms of the *definiendum*).[4]

At the end of the paper Quine turns to a definition of analyticity which is grounded in a verification theory of meaning. In this theory the meaning of a sentence is identified with its method of verification (or as a set of its verification/falsification conditions). An analytic truth can then be defined

4 In the first part of his argumentation, Quine considers a definition of analyticity according to which a statement is analytic iff it is derivable from logic and definitions alone. Of course, not all purported definitions can feature in such a derivation, only

as a statement that is confirmed come what may. This definition seems to satisfy at least the formal conditions for an adequate definition: the notions in the *definiens* are reasonably clear and the definition isn't circular. It would also allow us to define related semantic notions such as *synonymy*: two statements are synonymous iff the two statements have the same verification conditions.

However, the verification theory (as described) presupposes that there is a determinate set of verification conditions for individual statements. Quine argues that this presupposition is mistaken. Let us consider the verification conditions to be a set of *observation sentences*, i.e. statements about possible observations. It seems that which such observation sentences a given statement is associated with is not just a matter of that statement. To see this, take the simple example of an empirical test of a theory in a given experiment. Theories themselves do not directly imply any observation statements. They only imply them under further assumptions about what state the world is in (e.g. the experimental set-up) and what you are looking at or measuring. Let us take T to be our theory, B_{1-n} a set of background conditions that the theory requires in order to predict any specific event E and C_{1-n} a set of hypotheses, definitions and further conditions that specify what would be observable if such an event were to occur. Let us call a statement describing that observation O. Let us also assume that we observe $\neg O$. In this case, our "verification condition" does not simply and directly relate to T. It relates to T only via all these other statements. What we are dealing with is the following inconsistent set: $\{T, (T \rightarrow (B_{1-n} \rightarrow E)), B_{1-n}, (C_{1-n} \rightarrow (E \rightarrow O)), C_{1-n}, \neg O\}$. In order to deal with

correct ones, in which *definiens* and *definiendum* are really synonymous. Now, 'synonymous' is as problematic as 'analytic', but we might consider a definition of synonymy according to which two expressions are synonymous iff they can be replaced *salva veritate* for each other in all sentences of the language.

Whether that will give the right result, however, depends on the expressive resources of the language, in particular on whether the language contains intensional notions, such as 'it is necessarily the case that' (otherwise all coextensional expressions such as 'is an animal with a heart' and 'is an animal with kidneys' come out synonymous if all and only animals with kidneys also happen to have hearts). But what is the meaning of these intensional notions? Quine holds that the only somewhat intelligible analysis of 'it is necessarily the case that' is in terms of analyticity (a statement is necessarily true iff it is analytic). But now we came around in a full circle.

the recalcitrant observation, we have to do something with that set. But what we do is not forced by logic or observation. We have a choice now: we can either give up T or, in fact, give up any other statement from the set.

But since there is always such a choice, we can hold on to *any* statement come what may. On the other hand, we may chose to revise some of the definitions in the set C_{1-n} when accommodating $\neg O$. But such definitions were supposed to be analytic. Quine suggests that the possibility for empirical revision doesn't stop there. We may not only revise analytic truths when accommodating recalcitrant experience, we can even chose to revise a law of logic (in the example above, we could try to restrict logic such that O is no longer derivable). Thus the distinction between analytic and synthetic truths breaks down. Any statement can be held come what may, any statement is *a posteriori* revisable. But then there is no difference in principle between statements of logic and other statements when it comes to their empirical revisability.

This rejects the other two components of the standard account: logic can be revised, and there is no difference between empirical statements and the principles of logic, they are equally "*a posteriori*".[5]

All three dogmas of the standard view seem wrong. Logic is revisable, it is as *a posteriori* as everything else and it is unjustifiable. However, this package of views doesn't seems to be as stable (or convincing) as the standard view which held that logic is *a priori* certain and perfectly justified. The reason that the revised view seems less attractive is that, while Quine's result seems to show that statements of logic aren't special, Goodman's result seems to make them very different from our other beliefs. After all, it doesn't seem to be a feature of our ordinary empirical beliefs that they are unjustified, at least not for the reasons for which Goodman holds the principles of deduction to be unjustifiable. It seems we should look into the whole matter in more detail. We'll do this in reverse order. First we will reconsider logic's (rational) revisability, then have another look at the matter of logic's apriority and end the chapter with a brief discussion of the justifiability of logical principles.

[5] It's perhaps a matter of taste whether one wants to conclude from Quine's observations that the principles of logic are *a posteriori* or whether it is better to express Quine's point as establishing that the *a priori*/*a posteriori* distinction can't be made.

The Revisability of Logic

In order to systematize the discussion a bit, we will follow Graham Priest (2014) and distinguish between *logica docens*, *logica utens* and *logica ens*.

Logica docens is the "logic we teach", a theory that describes or models *logical consequence* and *logical truth*.

Logica utens is the "logic we use", the way we actually reason under idealized conditions. It refers to our competence, not our performance.

Logica ens is the "logic as it is" – what this refers to will depend on the metaphysics of logic we assume. If we are platonic realists, it will refer to certain relational facts that obtain between abstract objects; if we are structural realists, it will refer to certain very general facts; if we subscribe to some version of psychologism, it will refer to the *logica utens*; if we are non-cognitivists, it will refer to the *logica docens*.

For each of these three notions we can now ask whether logic is revisable, and, if it is, how it is revisable.

The Revisability of *Logica Ens*

We will begin with the *logica ens*, since that seems to be the easiest case. If we assume platonic or structural realism about logic, then the logical facts are independent of us – of how we actually reason as well as of what we believe about logic. If that's so, then we can't revise these facts either. This doesn't automatically follow from their mind-independence. For example, global warming is (as far as we know) a mind-independent phenomenon, a fact about the world, but one that we (hopefully) can influence and change. However, the way that realism thinks about the logical facts, they will be very fundamental facts or facts in a realm that is causally isolated from us. If that's so, we can't revise *logica ens*.[6]

[6] Graham Priest (2014) argues that *logica ens* can be revised, if the logical facts are relations between meanings (conceived as abstract objects). Although they are not changeable, we can change which meanings our words express (and thus *logica ens* can change after all). But then the change in logic is a change in the *logica utens*, and the idea that logic is "really" about the relations between the abstract objects is as confused as the view that physics is "really" about mathematical objects because we use mathematical objects to model physical phenomena.

On the other metaphysical conceptions, *logica ens* will reduce to *logica docens* or *logica utens*. So we should look at whether they can be revised, and, if they can, how they can be revised.

The Revisability of *Logica Docens*

The revisability of the *logica docens*, on the other hand, seems to be rather uncontroversial. As Graham Priest (2014) points out, logic changed considerably over the centuries. Moreover, the way it changed wasn't simply cumulative (in the sense that while we learned more about logic we never actually had to give up any of our beliefs). Some arguments that had been considered valid since Aristotle are now seen to be invalid. Priest gives as an example the syllogism *Darapti*:

$$\frac{\begin{array}{c}\text{All } A\text{s are } B\text{s}\\ \text{All } A\text{s are } C\text{s}\end{array}}{\text{Some } B\text{s are } C\text{s}}$$

We now believe this structure to be invalid, because it requires additional existence-assumptions (that there are any *A*s). But if *logica docens* has been revised, it obviously *can* be revised.

The process by which *logica docens* is revised has already been described above, when we described the *prima facie* epistemology of logic. We use judgements about the validity of particular inferences as well as general principles as input into a reflective equilibrium process which is further guided by general considerations of theory choice, such as simplicity, non-*ad hoc*ness, unifying power, fruitfulness, etc. If *logica docens* is revised in this way, it is certainly rationally revised. Depending on how we construe the subject matter of logic, we might be using sub-optimal data (when merely feeding intuitions and antecedently held beliefs into the process), but this doesn't make the process irrational.

The Revisability of *Logica Utens*

The revisability of *logica utens* is a less straightforward affair. Some philosophers considered the idea that we could revise the logic by which we reason to be so puzzling that they suspected it would lead to a paradox. Are the fundamental principles of logic not the premises of every argument by which we revise our beliefs? But, if so, wouldn't an argument that attacked one or several of these principles "saw off the limb on which the argument rests" (Katz, 1998, 73)?

Maybe that worry can be countered if the logical principles that occur in the revision are only principles that are themselves not subject to revision (cf. Colyvan, 2006; Field, 2004). Perhaps an account can be developed that can explain how a logical principle can feature in a form of reasoning that undermines itself. As Mark Colyvan (2006) points out, an argument that rests on premises that are undermined by that very argument is not paradoxical at all – all *reductio ad absurdum* proofs work like that.

Hartry Field identifies a second puzzle for revision of *logica utens*: a methodology that would advise us to follow another methodology would seem inconsistent. If we rationally change methodologies, then the rational motivation would have to come from a more fundamental methodology, which is not up for revision. Thus, if a method is rationally revisable, then it isn't fundamental. Hence, either logic isn't revisable, or logic is not, after all, our most fundamental methodology (Field, 2004, 2).

However, situations in which we might be led to a rational revision of logic don't seem to lead to such difficulties. Take, for example, the Liar Paradox (which we will discuss in more detail in Chapter 9):

(6.1) Sentence (6.1) is false.

With a standard theory of truth in a standard logic, sentence (6.1) creates a paradox. There are several ways to deal with that problem, but it seems unacceptable – by the lights of standard logic itself – just to accept the paradox. In this case, if we can convince ourselves that all other solutions to the paradox fail (or seem less attractive), we can consider revising standard logic (as in Field, 2008). Such a revision would be, in a way, recommended by our logic – *standard logic fails by its own standards.*

The Adoption Problem

But how should we just switch to another logic, another form of reasoning? Isn't *that* the absurd idea, that we could just rationally decide to change the very way we think? This problem has in recent years been discussed as the "adoption problem". Saul Kripke discussed it in seminars and lectures in the 1970s, but never published his view on the matter. Nevertheless, an academic discussion has evolved on the basis of these unpublished ideas.

Kripke argues that we can't just adopt a basic logical principle, and consequently can't adopt an alternative logic either. Kripke illustrates his point with the example of someone who doesn't yet see that each instance follows from a universal statement:

> Let's try to think of someone – and let's forget any questions about whether he can really understand the concept of "all" and so on – who somehow just doesn't see that from a universal statement each instance follows. *But* he is quite willing to accept my authority on these issues – at least, to try out or adopt or use provisionally any hypotheses that I give him. So I say to him, 'Consider the hypothesis that from each universal statement, each instance follows.' Now, previously to being told this, he believed it when I said that all ravens are black because I told him that too. But he was unable to infer that *this* raven, which is locked in a dark room, and he can't see it, is therefore black. And in fact, he doesn't see that that follows, or he doesn't see that that is actually true. So I say to him, 'Oh, you don't see that? Well, let me tell you, from every universal statement each instance follows.' He will say, 'Okay, yes. I believe you.' Now I say to him, "'All ravens are black" is a universal statement, and "This raven is black" is an instance. Yes?' 'Yes,' he agrees. So I say, 'Since all universal statements imply their instances, this particular universal statement, that all ravens are black, implies this particular instance.' He responds: 'Well, Hmm, I'm *not entirely sure*. I don't really think that I've got to accept *that*.' (Padro, 2015, n49)

In order to make his point, Kripke reminds us of Lewis Carroll's famous dialogue between Achilles and the tortoise, which we already encountered in the last chapter. There we discussed it as presenting an obstacle to conventionalism, the idea that the principles of logic are linguistic conventions. Kripke seems to think that the problem simply applies in the same way to the idea that logic could be rationally revised. As in the example above, it seems that you already have to be able to reason according to the rule of universal instantiation in order to know how to apply the rule. Likewise in the convention case: in order to know how to observe any conventions, you already need to know what the convention implies for a certain case, but that, in turn, already requires a logic.

But is the problem really the same? Let us grant for the moment that there is indeed a problem for conventionalism (in the last chapter we argued that there is reason to think there isn't). Then the argument shows that you can't get a logic out of thin air just by agreeing on a convention.

That was Quine's point: you already need a logic in order to know what the convention you agreed on implies.

The case of logic's rational revisability is not quite the same. Here we can assume that a logic is already in place. So, the argument only shows that for *some* rules, like perhaps universal instantiation, you already need some rule of that kind to adopt this rule, but why should this hold for *all* rules? For example, take disjunction introduction (as presented as a rule of the natural deduction system):

$$\frac{\phi}{\phi \vee \psi} \vee\text{-Intro}$$

Where does adopting this rule require this very rule? Also, it seems that the cases of revisions Quine has in mind are situations in which rules are *weakened*. Recalcitrant experience suggests that I revise my system of beliefs. Should I revise a theory, an observation statement, or the logic that tells me that the observation statement is in conflict with the theory? If I opt for the latter, I opt for weakening the logic that I already have. Is that difficult to do? Isn't that the same as when, after learning in a logic course that *Denying the Antecedent* is a frequent fallacy, you are less prone to make that fallacy?

So, the problem seems at best to be that *certain* basic logical principles can't be simply "adopted", but that doesn't speak against the idea of logic's rational revisability (because the latter might not at all concern the adoption of those particular principles).

Moreover, there seems to be a pragmatic solution to the adoption problem. Graham Priest (2014) argues that it is a matter of training, how we reason about certain subject matters. We might be trained in standard logic, and thus find reasoning this way very natural. But we might get training in another logic, and then follow that training if we have reason to believe that doing so is beneficial.

Is Logic *A Posteriori*?

So let's assume that logic – in any case *logica docens* and *logica utens* – can be revised. In the previous section we looked at one reason for the revision of logic: the Liar Paradox. This is a case in which our evidence against the logic is *a priori*. Could there be evidence against a logic that is *a posteriori*?

Let us first look at the matter from a metaphysical point of view. If we are non-cognitivists about logic, then it is a question of what we believe the logic-discourse is good for in order to decide which reasons we might have for changing that practice. But, whatever these reasons are, they won't be evidential. If we are platonist realists, it is also unlikely that there could be empirical reasons for revising our beliefs about abstract objects (but see Maddy (2014) for a naturalist epistemology for platonists). If, on the other hand, we subscribe to some form of psychologism, or structuralist realism, it doesn't seem absurd at all to suppose that there could be empirical reasons for changing our *logica docens* and/or *utens*.

Let us consider a psychologistic view first. Since psychologism holds that *logica docens* describes the actual ways in which we reason, empirical evidence about the ways we reason is clearly relevant for revising *logica docens*. But empirical evidence could also be relevant for revising *logica utens*. For example, we might find out that reasoning in some other way than the one we actually employ is psychologically beneficial – perhaps reasoning in a three-valued logic makes us more optimistic and enhances our well-being.

If we are structural realists, we in any case believe that *logica docens* describes the empirical world and not the way we reason about it. Thus, if we learn that our *logica docens* does not correspond to the empirical world, we should revise it. If that means that we are not reasoning according to the way the world is, we should revise *logica utens* too. And, indeed, Hilary Putnam (1975) argues that we should revise our logic in light of the puzzling results of quantum mechanics.

Non-factualist Epistemology and the Status of Logic

However, Hartry Field has argued that perhaps the question whether logic is *a priori* isn't after all a factual question, and there might be good reasons to treat logic as empirically indefeasible (Field, 1996, 1998). Field considers the following two necessary conditions to define the relevant notion of *a prioricity* for principles of logic:[7]

(A1) It is reasonable to believe the principles of logic without any empirical evidence for them.

[7] (A1) and (A2) are quoted from Field (1998, 1).

(A2) Logical principles are empirically indefeasible, in the sense that no possible combination of observations should count as evidence against them.

Field (1996) calls a principle that satisfies only (A1) "weakly *a priori*", and calls it "strongly *a priori*" if it also satisfies (A2). Whether principles of logic satisfy (A1) and (A2) is, according to Field, not a fully factual question. Surely, one can ask descriptively whether a principle satisfies (A1) and (A2) in a given *evidential system*, i.e. the rules that a cognitive agent is following that govern what he should believe in certain circumstances. But cognitive agents may have all kinds of weird evidential systems. So, the question of whether logical principles satisfy (A1) and (A2) isn't a question of whether there is some evidential system or other in which they do, but the evaluative question of whether or not it is a good thing to have an evaluative system that treats logical principles so that they satisfy (A1) and (A2).

It seems rather straightforward to argue that a decent evidential system should have logical principles satisfying (A1). As Field says,

> an evidential system that did not allow deductive reasoning until evidence favouring it was in would have nothing to recommend it, and it is hard to imagine what sort of evidence could ever be gathered by such a system that would allow deductive reasoning to begin when it was initially prohibited. (Field, 1996, 365)

But what about (A2)? Field argues that it is very hard to conceive of an epistemic system that would treat logical principles as failing to satisfy (A2). In contrast, it is relatively easy to imagine an evidential system that treats Euclidean geometry as empirically defeasible. It would just be a system that says (in some form or other) that we should give up Euclidean geometry on the basis of evidence E if someone comes up with a not-too-complex alternative geometry which in combination with highly plausible auxiliary hypotheses leads deductively to E, whereas no plausible auxiliaries can be found that together with Euclidean geometry lead deductively to E (Field, 1996, 369). Note that this evidential system for geometry mentions *deduction*. In fact, all evidential systems that we can imagine, including those that are inductive, will mention deduction. If so, then it is just very hard to see what a system should look like that treats deduction itself as empirically defeasible. But if it is hard to imagine such a system then it's certainly not plausible to have such a (at least as yet) inconceivable system evaluated higher than our actual evidential system which – as Field argues – does treat logical principles as satisfying (A2). Therefore – for all we know – we should be *a priorists* about the principles of logic.

At least this seems to establish that there is no alternative to an evidential system that treats *some logic or other* as strongly *a priori*. Field wants to establish a stronger claim: that an evidential system treats *classical* logic as *strongly a priori* (Field, 1996, 369, n10). Again, he argues that no non-classical logic is worked out in a way that would make clear what it would mean to use it as our everyday, all-purpose *logica utens*. As we have seen above, some logicians (e.g. Graham Priest (2014), but also Jodi Azzouni (2014) as well as Field's later self (Field, 2008)) would disagree with that claim.

The Justification of Logic

In order to understand the complexities of the issues involved, it is worth looking at the matter from the perspective of justification again. How are we justified in our fundamental logical beliefs? Following Paul Boghossian (2000), we will distinguish two issues here. The first is the question of how a cognitive agent *S* can be *justified* in the *belief* that, say, *modus ponens* is necessarily truth-preserving/valid. The second is the question of how *S* can be *entitled* in her *disposition* to reason according to *modus ponens*. Let's begin with the first of these. On first pass, there seem to be the options shown in Figure 6.1.

First candidates for a non-inferential justification would be that we can just "see" or "rationally intuit" that *modus ponens* is valid. But these notions are notoriously contested. A first candidate for an inferential defence of *modus ponens* would consist in some sort of argument that shows that *modus ponens* is valid, perhaps something like a soundness proof. But any such argument would – it seems – have to make use of *modus ponens* at some place or other. But then, as we have said above already, such an argument

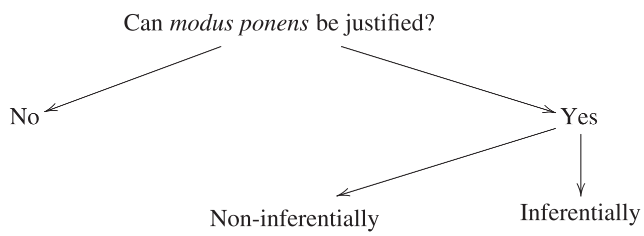

Figure 6.1 The justification of *modus ponens* (Boghossian, 2000, 230)

would be *rule-circular*. If that's unacceptable, we seem to be stuck with the result that Goodman and Hume arrived at: we can't justify our belief that *modus ponens* is valid.

But this seems unacceptable, for there is a plausible bridge principle between justification and entitlement (Boghossian, 2000, 234):

(6.2) If it is impossible for us to be justified in believing that a certain logical rule is truth-preserving, we cannot be entitled to reason in accordance with that rule.

But then, any belief that rests on deductive reasoning is unjustified. Since this also includes (6.2) itself, this looks like an extremely unattractive and unstable position. What to do? Can the conclusion be avoided? (In other words, is there a *straight* solution?) Can the conclusion be made more stable? (Is there a *sceptical* solution?) As Boghossian explains, there seem to be several options, represented in Figure 6.2.

Scepticism, on the branch at the far left, is the unstable position that we should avoid. 'NF about Logic', to the right of it, refers to non-factualism about logic. That's the kind of non-cognitivism about logic that we met in Chapter 5. On that account, apparent claims about logical truth, validity, etc. do not express beliefs in the first place, but express something else (commitments or perhaps inference licences; see Chapter 5 for details). This option then suffers from the problems of non-cognitivism about logic. But it would be an otherwise stable position: it's not a problem that we lack a justification for *modus ponens*; since there are no facts about logic in the first place, there is no such belief that would need to be justified.

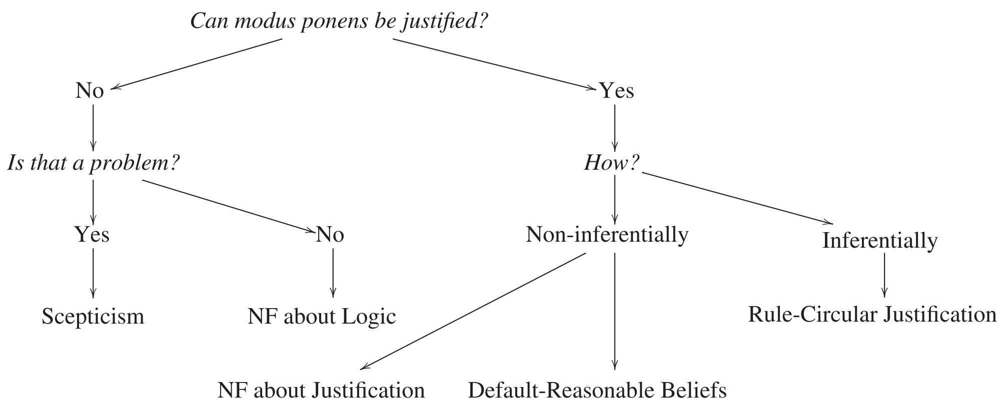

Figure 6.2 The justification of *modus ponens* continued (Boghossian, 2000, 236)

Let us turn to the options to non-inferentially justify *modus ponens.* Both candidates here are somewhat familiar from the previous subsection. 'NF about Justification' refers to non-factualism about justification. In a way this is similar to the view that Field was recommending (a partial non-factualism about epistemology). However, Boghossian has a stronger view in mind. This option here claims that – although judgements about logic (validity, implication, etc.) are factual – claims about what justifies what (or about which evidential system we should prefer) are not. Boghossian argues that this can't make sense of cases in which we are criticizing others for holding unjustified beliefs. Intuitively, there is something that does or doesn't ground our being right in doing so. But on this non-factualist view there is nothing to back us up.

The idea that, instead, *modus ponens* is non-inferentially justified, because it is a default-reasonable belief, should also be a familiar thought from the previous section. Boghossian argues that, although there are some default-reasonable beliefs, there must be a reason why a given belief should be considered default-reasonable, and he can't see why that should be the case for *modus ponens.* However, in the last section we met such a reason in Hartry Field's argument for endorsing weak *a priority* of logic: an evidential system that wouldn't treat *modus ponens* as default-reasonable would have nothing to recommend it – since you need deduction to get any kind of inferential justification off the ground, the principles of deduction should be endorsable prior to any evidence.

Boghossian's own sympathies lie with the branch at the far right of the tree, the rule-circular justification of *modus ponens*. This option is plagued by two difficulties. One, of course, is the circularity itself. Assume that you argue with someone who doubts *modus ponens.* Clearly, an argument that would use *modus ponens* is not going to win over someone who is initially sceptical about it. Relatedly, what does one's own entitlement in using *modus ponens* as a rule consist in, in an argument for its validity, if not in the belief that *modus ponens* is truth-preserving.

The second problem is known as the *bad-company* objection: if we allow that *modus ponens* can be justified in a rule-circular manner, then how can we block rule-circular justifications for all kinds of crazy rules? Just remember Arthur Prior's "tonk" from Chapter 4, the connective that had the introduction rule of disjunction and the elimination rule of conjunction. If rule-circular justifications were unconditionally allowed, we could justify reasoning with tonk as follows (Boghossian, 2000, 247):

1.	$\ulcorner P$ tonk $Q\urcorner$ is true iff $\ulcorner P\urcorner$ is true tonk $\ulcorner Q\urcorner$ is true	Meaning postulate
2.	P	Assumption
3.	$\ulcorner P\urcorner$ is true	2, T-scheme
4.	$\ulcorner P\urcorner$ is true tonk $\ulcorner Q\urcorner$ is true	3, tonk-introduction
5.	$\ulcorner P$ tonk $Q\urcorner$ is true	4, 1, biconditional-elimination
6.	P tonk Q	5, T-scheme
7.	If P, then P tonk Q	2–6, conditional-introduction

The last line of that argument is a canonical statement of tonk-introduction, which depends only on the meaning postulate in line 1. What's wrong with this "justification", if we allow rule-circular arguments?

Boghossian suggests addressing both of these problems with the help of an inferentialist conception of meaning for logical constants. We have already met such conceptions in Chapter 4. According to such an inferentialist conception, there is a set of inferences involving 'if… then…' (a subset of all the inferences that 'if… then…' can feature in) that are meaning constitutive for 'if… then…' for a thinker – 'if… then…' expresses the unique logical constant, if there is one, the semantic value of which makes the inferences in that subset truth-preserving (Boghossian, 2000, 249). Assuming that such an inferentialist account works, Boghossian proposes the following principle:

(L) If M is a genuinely meaning-constituting rule for S, then S is entitled to infer according to M, independently of having supplied an explicit justification for M.

As Boghossian argues, (L) is intuitively plausible. If certain inferences are indeed meaning-constituting for 'if…then…', then we couldn't even have the belief that we are entitled to reason according to those rules, without having the disposition to do so, prior to and independently of an explicit justification of those rules.

Boghossian also argues that such an inferentialist account can address the bad-company objection. (L) supports the following restriction on rule-circular justifications:

(RC) *S*'s rule-circular argument for a rule of inference *M* will confer warrant on *S*'s belief that *M* is truth-preserving, provided that *M* is a genuinely meaning-constituting rule for *S*.

With (RC) we can block the argument for tonk provided above. The rule-circular argument does not confer warrant on its conclusion, because the tonk-introduction rule is not meaning constitutive. We can see this by observing that we can't construe a truth-table for tonk from its introduction and elimination rules.

Thus, if inferentialism for logical constants is plausible, then we can address the bad-company objection and the intrapersonal circularity worry. The view doesn't help with the second half of the circularity problem: how should we convince somebody antecedently sceptical about *modus ponens* about its truth-preserving qualities? Boghossian accepts this and distinguishes *suasive* and *non-suasive* reasons. We have non-suasive reasons for endorsing *modus ponens*, which – unfortunately – don't have the power to move a sceptic.

Of course, the fact that the view depends on inferentialism about the meaning of logical constants is "a big 'if'". Timothy Williamson (2007) provides powerful arguments against the view that the meaning of logical constants in the mind or the utterances of a cognitive agent is determined by any set of inferences that that cognitive agent would accept as valid.[8] According to Williamson, the meaning of these connectives is fixed by the public language. But if that's so, then (L) loses its plausibility.

Williamson asks us to consider Peter and Stephen, both native speakers of English. Peter believes that

(6.3) All vixens are vixens.

not merely presupposes, but entails, that there is at least one vixen. This is because he, in general, came to believe that universal quantification is existentially committing. Moreover, Peter also happens to believe that there are actually no vixens, since he has been convinced that foxes are merely an invention of MI6, who planted all the evidence for the existence of foxes in order to involve people in protests about fox-hunting rather than in

[8] These arguments will be familiar to you from an adaptation that we considered in Chapter 4.

Table 6.1 *Kleene table for the conditional*

$\rightarrow$	1	0	#
1	1	0	#
0	1	1	1
#	1	#	#

protests about the war in Iraq (Williamson, 2007, 86–87). Consequently, Peter denies (6.3) and doesn't assent to it.

Stephen, on the other hand, believes that borderline cases for vague terms lead to truth-value gaps. Stephen treats truth-value gaps as the third value in Kleene's "strong tables". The Kleene table for the conditional is shown in Table 6.1.

Stephen also believes that some female evolutionary ancestors of foxes are borderline cases of 'fox', and hence borderline cases of 'vixen'. But then, for such an animal as the value of 'x', 'x is a vixen' is neither true nor false, thus 'x is a vixen $\rightarrow$ x is a vixen' is neither true nor false by the table given.

For Stephen, a universal generalization, such as '$\forall x(Fx \rightarrow Gx)$', is true if '$Fx \rightarrow Gx$' is true for every value of 'x', and false if '$Fx \rightarrow Gx$' is false for some value of 'x'. Since 'x is a vixen' is not true for every value of 'x', nor false for any value of 'x', 'Every vixen is a vixen' is neither true nor false, according to Stephen (Williamson, 2007, 87–88).

Peter and Stephen are peculiar in the inferences they accept as valid and the statements they accept as logically true. But, so goes Williamson's argument, there is little room to hold that in Peter's and Stephen's use the logical constants just acquire some deviant meaning:

> Peter and Stephen are native speakers who learned English in the normal way. They acquired their non-standard views as adults. At least before that, nothing in their use of English suggested semantic deviation. Surely they understood [6.3] and its constituent words and modes of construction with their ordinary meanings then. But the process by which they acquired their eccentricities did not involve forgetting their previous semantic understanding. [. . .] By ordinary standards, Peter and Stephen understand [6.3] perfectly well. Although their rejection of [6.3] might on first acquaintance give an observer a defeasible reason to deny that they understood it, any such reason is defeated by closer observation of them. (Williamson, 2007, 90–91)

If Williamson is right, then inferentialism can't provide a satisfactory epistemology for logic either (which might make the non-inferential justifications of basic logical principles look more attractive again).

Questions

1. Above we said that the *adoption problem* might not arise if our revision requires weakening the logic that we have. For which cases of strengthening the logic does it arise then? (For all cases of strengthening? Only certain ones?)
2. Continuing the topic of the previous question: what reasons could there be for suggesting a strengthening of a logic?
3. Williamson's argument doesn't seem to establish that there is something wrong with inferentialism as such but only with an individualistic version of the view. (Peter's and Stephen's individual dispositions to use logical constants did not impact the meaning of those constants.) Does that create a loophole for Boghossian's account? Can a non-individualist version of inferentialism combine with Boghossian's argument?

7 Logical Pluralism

Logical pluralism is, roughly, the view that there is more than one correct logic or, alternatively, that there is more than one genuine consequence relation, more than one right answer to the question of whether and why a given argument is valid, more than one collection of valid inferences (or of logical truths), or more than one right way of reasoning. But these rough characterizations are already a sample of how many different versions of the thesis of logical pluralism there can be, corresponding to the different ways in which one can specify more carefully what a *logic* is, and what it would be for a logic to be *correct*.

Those different versions of the thesis of logical pluralism range from the nearly innocuous to the highly contentious. Using the terminology of previous chapters, take 'logic' to mean *logic as a theory*, especially a *pure* one. In that sense, a logic is a mathematical theory. Now take 'to be correct' to mean the satisfaction of standards of correction for mathematical theories, say, self-consistency, among others. Then the claim 'there is more than one correct logic' just means 'there is more than one pure logic(al theory)', which is nearly anodyne now. We say "nearly anodyne" because this idea is relatively recent. To take just one example, Quine (1970) claimed that many non-standard logics are not really logics but algebras. Even the mere idea of "pure logic" can be incoherent for some theorists, since a logic is for them essentially interpreted in a particular way, namely as dealing with the evaluation of linguistic-like entities, as we saw in Chapter 1.

Without denying that interesting and substantial debates might arise, and indeed have arisen, around these other senses of 'logic', we will in this chapter stick to the canonical side of the issue. Thus, we take 'logic' to mean *logic as a canonically applied theory*, i.e. a theory about the canonical logical phenomena: right reasoning in sciences or ordinary language. Additionally, we take 'to be correct' for a logic in this sense to mean that the logic is "materially adequate" a la Tarski, that is, it has all the relevant features that logical consequence intuitively and informally has in ordinary language. Then, 'there is more than one correct logic' would mean that *there is more than one theory that has all the relevant features that right reasoning has in ordinary language*. More briefly, in this reading, 'there is more than one correct logic' is spelled out as 'there is more than one correct theory as to what constitutes right reasoning'. This is certainly not anodyne. How could it be that a theory implying that a statement of the form $A \vee \neg A$ can be deduced from any other statement whatsoever, *and* a theory implying that this is not the case, are *both* correct: that both represent all the relevant features of right reasoning in ordinary language? These theories seem to plainly contradict each other, and not be supplementary in any sense.

A Plurality of Pluralisms

Suppose that the enterprise of finding one or more correct logics for reasoning in ordinary language is a bit ambitious, especially because it may be quite indeterminate what "correct reasoning in ordinary language" is supposed to include. So let us abstract a little bit and focus on *domains of discourse*. Of course, right reasoning in ordinary language can itself be a domain of discourse, perhaps an all-encompassing domain of discourse, but typically domains of discourse will be more restricted, for example to (parts of) mathematical activity, or (parts of) the quantum domain or other domains of scientific enquiry. Then we can distinguish the following pluralist claims:

Pointed logical pluralism: There is exactly one domain of logical enquiry for which there is more than one correct logic.

Collected local logical pluralism: There is a domain for which there is more than one correct logic.

Distributed local logical pluralism: There is more than one logic that is correct for some domain.

Universal local logical pluralism: For all domains there is more than one correct logic.

Collected local logical pluralism is stronger than *distributed local logical pluralism*. The first says that there are at least two distinct logics L_1 and L_2 and a domain D such that L_1 and L_2 are both correct for D. The second says something weaker, namely that there are at least two distinct logics L_1 and L_2 and at least two domains D_m and D_n such that L_1 is correct for D_m and L_2 is correct for D_n. But even the weaker form of local logical pluralism is contentious. A theorist might reject the idea that logic is not applied globally but to a restricted domain. Accordingly, there are several versions of logical monism:

Pointed logical monism: There is exactly one domain of logical enquiry for which there is exactly one correct logic.

Collected local logical monism: For some domains there is exactly one correct logic.

Universal local logical monism: For all domains there is exactly one correct logic.

Universal logical monism: There is exactly one logic that is correct for all domains.

These formulations of logical monism and logical pluralism allow more or less easily the recognition of other options in the debate. For example, the idea that there is *no* logic that is correct for the single domain of logical enquiry, or the idea that *every* logic is correct for *some* domain of enquiry. Options similar to them have appeared very recently under the labels of *logical nihilism* and *logical universalism*, respectively. Nonetheless, they are options allowed not only by the combinatorics of the quantifiers, but also by some internal problems of logical pluralism and logical monism. In the next sections we will explore some of the usual defences of different versions of logical pluralism.

Carnap's Tolerance

Carnap's *logical tolerance* is usually regarded as one of the earliest forms of logical pluralism. In Carnap's view, logic(s) is (are) concerned with a single area, namely the evaluation of arguments, but arguments are always arguments-in-a-language and there are many languages; there is a single set of standards of evaluation but the features proper of each language will yield different valid (and invalid) arguments. The single area of evaluation of arguments can be decomposed into many domains of enquiry: evaluation of arguments in physics, in biology, etc. One can even associate different logics/languages to the same domain of enquiry (microphysics, mathematics, economics, etc.). Carnap's pluralism is based on his well-known *Principle of Tolerance*:

> *Principle of Tolerance: It is not our business to set up prohibitions, but to arrive at conclusions.* [. . .]
> *In logic there are no morals.* Everyone is at liberty to build his own logic, i.e. his own form of language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments. (Carnap, 1937, §17)

Carnap formulates this principle in his *Logische Syntax der Sprache* (*The Logical Syntax of Language*: Carnap, 1937). In this book he develops a syntactic (and also the first step towards a semantic) approach to studying languages and their logical properties. For Carnap, we can construe languages in different ways with logics of different strengths. It is only a matter of pragmatic choice what we want to consider "logical" and what we want to treat as "extra-logical". If we think that's useful, we may build all kinds of mathematical or even physical content into the language, depending on the purpose of application (as said above).

Since Carnap considers this to be a pragmatic choice, it seems clear that for him the question of whether an argument is valid must be treated as an *internal* question (Carnap, 1956). Only within a chosen language (containing a certain logic) does it make sense to ask whether an argument or an inference is valid. Whether it is valid then does depend on the rules of that language. An external question about an argument's validity, on the other hand, is meaningless. It can only make sense as an elliptic pragmatic question (should we adopt a language with such and such rules?). Stewart Shapiro has argued that the idea of tolerance (i.e. not to dismiss deviant logics out of hand) does not itself imply relativism or pluralism about logic. One might be tolerant towards other logics and openly consider their advantages while still believing that there is only one true logic out there (Shapiro, 2014b). However, it seems clear that this was not Carnap's view. If *everyone is at liberty to build his own logic*, then there is obviously more than one correct way.

There is a nearby form of logical pluralism, which is championed by Achille Varzi and presented as a form of *logical relativism*. As we have seen in Chapter 4, some – including Alfred Tarski at some point – hold that the demarcation between the logical and the non-logical expressions in a language is matter of pragmatic choice. But since logical consequence to some extent hinges on the choice of what we keep constant and what we allow reinterpretations for, logical consequence is relative to our choice of logical constants (Varzi, 2002).

It is a consequence of Carnap's view that his pluralism is confined to languages, thus *any difference in logical consequence is due to a difference in language* (Restall, 2002). Moreover, the formal languages, about which Carnap is a pluralist, are indeed those that we can adopt as *regimentations* of ordinary language.

Modelling Pluralism

This is different in Cook–Shapiro modelling pluralism. According to this version of logical pluralism, logic(s) is (are) concerned with a single area, namely *modelling* natural language, but models are always models-of-only-some-aspects; there is no single set of standards of evaluation of models – evaluation of models is always evaluation-according-to-some-purpose.

Consider for example a situation in which you want to explain the location of your house relative to some other location that is, say, three blocks away (imagine you're explaining to a friend how to get to your house from the bus stop). Perhaps you do that by drawing a crude map in the sand on the ground and model your house by putting a small stone on that map, and another stone that represents the bus stop. The stone is in many ways distinct from your house (fortunately), but it can serve as a model for your house in that situation. Now think of another purpose; perhaps an architect wants to discuss how a carport could be added to your house and so makes a small-scale model of it from cardboard. Again, your actual house doesn't have walls out of cardboard, but for the purpose at hand, this doesn't matter. The stone would not have been a good model now. It doesn't properly represent the shape of the house, and it wouldn't have helped the architect in communicating how a carport could be added to it in an aesthetically pleasing way, say.

Note that not only can different models serve different purposes, but there is also no "super-model" that could serve all purposes. Even an exact replica of your house, or even your house itself wouldn't do, since it is often essential for the purpose of a model that it abstracts away from aspects that the actual thing has. Take the first case. A “life”-sized house might do a poor job in representing the location of your house vis-a-vis some other location, since it might be too large to help your friend get an overview of the relative location. Thus even the actual thing is – for certain applications – a poor model of itself. So models are good or bad depending on purpose, and for different purposes there might be different models.

This form of logical pluralism is hardly controversial. But it leaves room for a more interesting kind of pluralism:

> Once we realize that logics are merely models of the consequence relation in natural language, we might wonder whether there could be two logics, such that (i) the logics belong to the same language and involve the same interpretation of the logical/nonlogical divide; (ii) the logics are incompatible – that is, they validate different sets of inferences; (iii) the logics are, all things considered, equally good models of logical consequence in natural language; and (iv) there is no third logic such that this logic is a better model of logical consequence than the two competing logics. (Cook, 2010, 501)

Shapiro (2014a) goes some way in establishing these four points for his “eclectic pluralism”. He argues that there are perfectly worthwhile mathematical structures, such as, for example, Kock–Lawvere’s infinitesimal analysis, that require to be studied in non-classical logic, since they are classically inconsistent. Shapiro argues that – even if there is sometimes a classical workaround to studying these structures, they are better understood in terms of their underlying logic than via some translation into classical logic. But then validity and consequence become relative notions – relative to mathematical structures. The underlying logics of these structures are incompatible, arguably belong to the same language and preserve the meaning of connectives, and they all serve an equally important purpose (they allow us to study worthwhile mathematical structures).

This view is motivated via the mathematical practice of studying all kinds of structures; it is a logical pluralism “for the working mathematician” (Hjortland, 2016). A similarly language-internal pluralism can be motivated on considerations of the vagueness of the logical consequence relation alone, as we shall see in the next section.

Beall/Restall Pluralism

Beall and Restall's (2000, 2001, 2006) *case-based logical pluralism* is a version of logical pluralism which has received considerable attention in recent years. In this version, logic(s) is (are) concerned with a single area, namely the evaluation of arguments; it is the same (form of) argument in the same language which is evaluated in different *cases*; cases provide different standards for such evaluation and they may yield different valid (and invalid) arguments. Thus unlike Carnap's view the pluralism arises within one language, and unlike the Shapiro/Cook view it is a pluralism about the logical consequence relation (and not one about different models thereof).

Beall and Restall motivate their pluralism by observing that there is an accepted explication of logical consequence in semantic terms, which is largely due to the work of Tarski and formulated in terms of quantification over cases.

Generalized Tarski Thesis (GTT) An argument is valid$_x$ if and only if in every case$_x$ in which the premises are true, so is the conclusion.

An important part of Beall and Restall's pluralism is that GTT is *unsettled*, that is, that no particular logic is obtained unless one provides cases, and there are at least two kinds of cases. If we consider (complete) possible worlds as cases, we obtain the consequence relation of classical logic; if we takes stages in a construction as cases, we obtain intuitionistic logic, and taking (potentially inconsistent) situations as cases, we end up with paraconsistent logic.

But there is more to Beall and Restall's pluralism than GTT and its unsettledness. Based on a more or less widespread understanding of logic, they have put forward an argument to exclude from the realm of "proper" logic what they take to be merely so-called "logics". They say:

> Logic, whatever it is, must be a tool useful for the analysis of the inferential relationships between premises and conclusions expressed in arguments we actually employ. If a discipline does not manage this much, it cannot be *logic* in its traditional sense. (Beall and Restall, 2006, p. 8, italics in the original)
>
> [. . .] any settling of the relation of logical consequence must be a necessary, normative and formal relation on propositions. (Beall and Restall, 2006, p. 29)
>
> We hold that deductive validity is a matter of the preservation of truth in all cases [. . .]. This analysis of validity owes a great deal to the tradition [. . .]

> It is also connected intimately with the constraints on consequence that we have already seen. [. . .] So, the analysis of logical consequence as preservation of truth in all cases goes some way to explaining how a relation of logical consequence is necessary, normative and formal. (Beall and Restall, 2006, p. 23f)

> Non-transitive or non-reflexive systems of 'entailment' may well model interesting phenomena, but they are not accounts of *logical consequence*. One must draw the line somewhere and, pending further argument, we (defeasibly) draw it where we have. We require transitivity and reflexivity in logical consequence. (Beall and Restall, 2006, p. 91, italics in the original)

The argument can be reconstructed as follows:

(BR0) A logic is an account of the relation of logical consequence (or validity).

(BR1) An account of logical consequence must be a tool useful for the analysis of the inferential relationships between premises and conclusions expressed in arguments we actually employ.

(BR2) Logical consequence has at least the features of necessity (the truth of the premises in a valid argument necessitates the truth of the conclusions), formality (valid arguments are so in virtue of their logical form) and normativity (rejecting a valid argument is irrational).

(BR3) Logical consequence is forwards truth-preserving (i.e. logical consequence preserves truth from premises to conclusions).

(BR4) Forwards truth-preservation is a (non-empty)[1] reflexive and transitive relation.

(BR5) Hence, logical consequence is a reflexive and transitive relation.

(BR6) Then, if a relation between the elements of a structure is either non-reflexive or non-transitive, such a relation is not one of logical consequence.

(BR7) Therefore, if no relation on a structure is logical consequence, such a structure is not a logic.

[1] In what follows we will always deal with candidates for logical consequence that are non-empty relations, so we will omit this qualification. Whether the empty relation can be rightly regarded as logical consequence would require a separate discussion.

One should not get distracted by the fact that Beall and Restall *defeasibly* draw the line at reflexivity and transitivity, as if it were not an important part of their philosophical view. As the argument shows, they base it on some crucial philosophical motivations and reasons. Of course, they may be willing to drop these requirements, but *pending further argument* they keep them. Below we will sketch an argument for actually dropping those conditions in terms shared by Beall and Restall and, in general, by those who accept this "core tradition" as a necessary part of logic.

The argument BR1–BR7 is relatively simple yet clarifies the kind of philosophical background that is operating when we decide what to accept as a logic, in this case what Beall and Restall call the "core tradition" of logic, namely that logical consequence is a necessary, formal and normative truth-preserving relation used to evaluate arguments we actually employ. This is important because their pluralism aims to be an adequate framework in which to understand contemporary logic (see Beall and Restall, 2006).

However, in defending the fact that logical consequence is forwards truth-preserving, and hence reflexive and transitive, they rule out some activity in contemporary logic, not only activity coming from mathematics and computer science which could be suspected of philosophical naivety, but also some philosophically motivated systems. For example, a suitable logic for so-called "epistemic gain" (Tennant, 1994), a logic for unrestricted set comprehension (Weir, 1998), a theory for an unrestricted truth predicate (Ripley, 2012), or a logic for vague expressions (Zardini, 2008), and even probably Aristotle's own non-reflexive notion of deduction, are all non-transitive.

Though nothing is completely uncontentious, we can assume for the sake of argument that there is hardly any disagreement about what Beall and Restall call "the core tradition" in logic, namely (BR0), (BR1) and (BR2); we do not count disagreements about how to spell out some crucial notions in the above – like formality or normativity, to mention just two – as disagreements with the spirit of these premises.[2] (BR3) has also been widely accepted.

[2] As has been noted by Paseau (2007), there might be disagreement on whether normativity, formality and necessity are the *only* settled features of logical consequence – others might be, say, aprioricity or universality, mentioned in Chapter 1 – but not that they are features of logical consequence. In any case, as we

(BR4) has been proved for certain formal languages, among them the most widely used in logic (for example, in Hardegree, 2005). (BR5) follows from (BR3) and (BR4) by the transitivity of the predicative 'is'; (BR6) from (BR5) by contraposition, and (BR7) by a generalization from (BR0) and (BR6).

Read (2003) has argued that (BR5) and (BR4) can fail, but not exactly by rejecting any of the premises, but by making (BR3) more precise and taking into account the role that relevance plays in truth-preservation. Recently there has been opposition to (BR3) in the following terms:

- Logical consequence is not forwards truth-preservation because there is nothing like (forwards) truth-preservation in all cases, on pain of triviality. (See for example Beall, 2009; Field, 2009b; Murzi, 2014.)

Another way of disputing (BR3) and (BR5) is this:

- Logical consequence is not forwards truth-preservation because there are other ways of logically relating premises and conclusions which are not reflexive or not transitive but count nonetheless as accounts of logical consequence as in (BR0), (BR1) and (BR2).

We will consider criticisms of the idea that logical consequence is a reflexive, transitive, truth-preserving relation in Chapter 9 in more detail. If one takes these considerations on board, one might well end up with a wider logical pluralism.

Be that as it may, the argument against non-reflexive and non-transitive logics (BR1–BR7) shows that it is not true that "Beall and Restall's choice of what can be dealt with pluralistically is idiosyncratic and not principled" (Bremer, 2014). Also, it is not the case that Beall and Restall incur in addition an "exclusion of logics that fail to meet monotonicity": they only require reflexivity and transitivity, not monotonicity, and indeed monotonicity is nowhere mentioned in the book. Moreover, arguments like those of Bueno and Shalkowski (2009) and Bremer (2014) for the idea that Beall and Restall's logical pluralism implies either logical universalism or logical nihilism fall short because they ignore these other essential aspects of Beall and Restall's proposal. Nonetheless, it will be instructive to consider such arguments for the exploration of other positions in the vicinity of logical pluralism.

---

have reconstructed the argument, it is not necessary to assume that these are the only ones.

According to these critics, Beall and Restall's logical pluralism, or rather something very similar to it, by the previous considerations, collapses either in *logical universalism* or in *logical nihilism*. Logical universalism is the thesis that every relation between premises and conclusions is a legitimate case of logical consequence. Given that the unsettledness of GTT is sufficient for logical pluralism, virtually every relation between premises and conclusions is a case of logical consequence; one only needs to construct more or less carefully the kind of cases in which GTT would be satisfied. Then logical universalism will follow. On the other hand, if one reads 'every case' strictly, then virtually no relation can count as logical consequence, as virtually every argument has a counterexample in some case. Then logical nihilism, the thesis that no relation can be a legitimate case of logical consequence, follows.

Quoting in full Bueno and Shalkowski's argument for the collapse of a cases-based logical pluralism into logical nihilism will be useful to illustrate why such logical pluralism is not Beall and Restall's:

> Only a very weak consequence relation survives this scrutiny, according to their accounting of the necessity constraint as quantification over all cases. This relation is just the intersection of the inferences treated as valid by classical, constructive, and paraconsistent logics. Some fragments of positive logic and some rules of identity will survive. This survival, however, is merely an artefact of having considered only the cases appropriate for the semantic underpinnings of classical, constructive, and paraconsistent logics. If expanding the domain of our metatheoretic quantification to accommodate the semantic underpinnings of constructive and paraconsistent logics is in order, so is the similar expansion to accommodate the semantic underpinnings of nonadjunctive logics, which would invalidate instances of conjunction introduction, and of certain quantum logics, which would invalidate typical laws of identity and distribution. Once the partisan spirit of logical monism is replaced with the open-minded embrace of cases suitable to alternative logics, no commonly promulgated consequence relation seems to satisfy the necessity constraint. Hence, according to their own accounting of the constraints on relations of logical consequence, there are no such relations, much less multiple relations of consequence. Their account of the necessity constraint on logics ends not in logical pluralism, but in logical nihilism. (Bueno and Shalkowski, 2009, 300)

The problem is that Beall and Restall's "open-mindedness of cases suitable to alternative logics" is not as open as to admit that the "expansions of the domain of our metatheoretic quantification to accommodate the semantic underpinnings of constructive and paraconsistent logics" is similar enough to the expansion required "to accommodate the semantic underpinnings of nonadjunctive logics, which would invalidate instances of conjunction introduction, and of certain quantum logics, which would invalidate typical laws of identity and distribution". And the reason is that Beall and Restall's logical pluralism is more than the open-mindedness derived from the unsettledness of GTT, but – as we have seen above – also includes a defence of reflexivity and transitivity of logical consequence via the idea that logical consequence is a species of truth-preservation. Bueno and Shalkowski's logics cannot be accommodated if they fail to be reflexive or transitive, as many quantum logics fail to be, because they would be non-truth-preserving relations and then would be modelling other interesting phenomena, but not logical consequence. This can be disputed (as we have shown above), but the collapses into universalism or nihilism are far from having been established.

Logical Monism

There is a simple argument for *universal logical monism*, purported to follow from the pretheoretical notion of logical validity, put forward by Graham Priest (2006), among others. Priest works with a broad notion of logic in the sense that he is ready to accept that inferential tools for certain particular domains augmented with principles specific to those domains count as logics, but he says that there is nonetheless *one true logic*, a logic whose inferences are valid in all domains and that lacks principles depending on specific domains. Priest presents his argument succinctly as follows:

> Is the same logical theory to be applied in all domains, or do different domains require different logics? [...]
>
> Even if modes of legitimate inference do vary from domain to domain, there must be a common core determined by the syntactic intersection of all these. In virtue of the tradition of logic as being domain-neutral, this has good reason to be called *the* correct logic. But if this claim is rejected, even the localist must recognise the significance of this core. Despite the fact that there are relatively independent domains about which we reason,

> given any two domains, it is always possible that we may be required to reason *across* domains [...] (Priest, 2006, 174f)

We hereby present a version of the argument using valid inferences, but it can be easily turned into an argument about logical truths. Read '$X \vDash Y$' as "Y is a logical consequence of X". We also use the word 'case', but it can be replaced by the word 'model' according to the preferences of the reader.

P0 An argument $X \vDash Y$ holds in a case if and only if, in that case, if X is true then Y is true. (The pretheoretical notion of holding in a case)

P1 An argument $X \vDash Y$ is valid if and only if it holds in all cases. (The pretheoretical notion of logical validity)

P1′ $X \vDash Y$ is not valid if and only if it does not hold in all cases. (From P1, contraposition)

P2 There is at least one collection of (enough) arguments holding in all cases. (Existence of a logic)

P3 There are at least two different collections of all arguments holding in all cases. (Logical pluralism, hypothesis to be reduced)

P4 If two collections of all arguments holding in all cases are different, then there is at least one argument $X \vDash Y$ such that it belongs to a collection but not to the other. (Extensionality of collections)

C1 Since they are different collections of valid arguments, there is an argument $X \vDash Y$ belonging to one of the collections but not to the other. (From P3, P4 *Modus ponens*)

C2 If $X \vDash Y$ is a valid argument, it holds in all cases. (From P1)

C3 If $X \vDash Y$ is not a valid argument, it does not hold in all cases. (From P1')

C4 $X \vDash Y$ holds in all cases. (From C1, C2 *modus ponens*)

C5 $X \vDash Y$ does not hold in all cases. (From C1, C3 *modus ponens*)

C6 $X \vDash Y$ holds in all cases and $X \vDash Y$ does not hold in all cases. (From C4, C5 adjunction)

C7 There are not even two collections of arguments which are different and hold in all cases. (From C6, *reductio* of a semantically untenable contradiction)

C8 There is exactly one collection of arguments holding in all cases. (Logical monism, from P2 and C7)

Logical pluralists might contend the argument claiming that the quantification "all cases (domains)" is not absolute, but should be read as "all cases (domains) *of a certain kind*". For example, standard first-order logic would stem from taking cases or domains to be consistent and complete possible worlds, whereas constructive logics would be given when cases are taken to be possibly incomplete bodies of information or warrants or constructions, while relevance logic would be given when cases are taken to be possibly incomplete or inconsistent (or both) ways the world might or might not be. Thus, there could be different collections of arguments logically valid in all cases, for they could be valid in all cases but of different kinds.

This pluralist reply seems not to be a good one, for then 'all the cases' does not mean "all the cases" and there is the risk of making logic dependent on the content or particularities of the case under consideration, which goes against the generality and topic-neutrality expected from logic. Of course, the pluralist might say that logic is not topic-neutral or formal etc., but a further argument for this is needed. Moreover, there is the possibility that the arguments valid in all the (different kinds of) cases would be regarded as the real logically valid arguments, for they are indeed valid in all cases, do not vary from case to case, and hence hold independently of the particularities of each case.

Logical Nihilism Enters the Scene

Another pluralist option, not widely studied yet, is to bite the bullet, to take the pretheoretical notion of logical validity at face value and then try to show that it might be inapplicable. As we have seen before, Bueno thought this is one of the natural consequences of Beall and Restall's logical pluralism. The logical monist assumes that the collection of valid arguments, defined as arguments holding in all cases, is not empty. We have seen in the preceding paragraph that a logical monist might insist on the existence of one true logic, claiming that the arguments valid across all the cases of every kind are the real valid arguments. This move rests on premise P2 of the previous section. But what if it were false, i.e. what if there were no arguments holding in all cases (of all kinds)? Would there be no logic at all? Some arguments by possibilists (see Mortensen, 1989; Estrada-González, 2011) and trivialists (Mortensen, 2005; Estrada-González, 2012) seem to imply that there are no inferences holding in all cases, and that is why authors like Gillian Russell (2016) have called such a position 'logical nihilism'.

However, the idea that there are no inferences holding in all cases hardly entails the nonexistence of any logic at all, unless one has bought the idea that a logic has to be valid in all cases to be a real logic. Even though there were no inferences valid in all of them, cases might need special inferences as inferential patterns ruling right reasoning in them. Premise P2 further requires a "large enough" number of valid inferences, for even if the collection of valid inferences were not empty, if it consisted of, say, only one or just a few inferences, it would be vacuous in practice to call such a small number of valid inferences a "logic". It seems then that logic would be better characterized as an inferential device and the universal quantifier in the definition of validity should be explicitly restricted:

(7.1) An argument $X \vDash Y$ is *k*-valid if and only if it holds in all *k*-cases.

As it is, this notion of validity is compatible with both the existence of one true logic (since it allows for the non-emptiness of the case of all cases) and the idea that logics may be inferential devices for specific domains, but both options require further argument. Priest rejects the idea that, in practice, every principle of inference – or at least a large number of them so as to make speaking of a logic vacuous – fails in some situations. His argument for this, premise P2, is that, to the extent that the meanings of connectives are fixed, there are some principles that cannot fail. However, this is not completely right.

Pluralism about Validity

The pluralist replies considered hitherto tried to provide a special account of the phrase 'all cases (or domains)' or attempted to give reasons to reject premise P2 in Priest's argument for *universal logical monism*. There is an additional way of challenging logical monism, not necessarily incompatible with the former and just recently being taken into account in the specialized literature. It consists of challenging premises P0 and P1, i.e.

challenging at least the uniqueness of the pre-theoretical notions of *holding in a case* and *validity*. For example, the following characterizations of validity

V1 An argument $X \vDash Y$ is valid if and only if in all cases in which X is true Y is true too.

V2 An argument $X \vDash Y$ is valid if and only if in all cases in which X is not false Y is true.

V3 An argument $X \vDash Y$ is valid if and only if in all cases in which X is true Y is not false.

turn out to be equivalent in classical logic, where truth and falsity are collectively exhaustive and mutually exclusive values. But when this is not the case, these different notions of validity may give rise to different collections of valid arguments, and hence to a plurality of logics with very different properties. Moreover, they can be used to show, *contra* Priest, how some principles can fail without changing the meaning of connectives.

The basic idea is as follows: suppose that the satisfiability conditions of a conjunction are the usual ones (omitting indices for simplicity): $\mathfrak{v}(A \wedge B) = \inf(\mathfrak{v}(A), \mathfrak{v}(B))$, with respect to an ordering between at least two truth-values, true (denoted '$\top$') and false (denoted '$\bot$'). One could say that its satisfiability condition constitutes the meaning of conjunction (under a model-theoretic perspective, at least). The only supposition about truth-values in order to give the satisfiability condition is that there is at least one of them (at least two if one supposes further that $\top \neq \bot$), but nothing is said on exactly how many of them there are.

Let us take a closer look at the notions of logical validity at work, especially V2. Suppose there are three truth-values, $\top$, μ and $\bot$, with the order $\bot < \mu < \top$. Take $\top$ as the value *true*, $\bot$ as the value *false* and μ as neither true nor false. An argument like $A \wedge B \vDash A$ fails using V2 even though the meaning of conjunction, its satisfiability condition, is the usual one. Take for example $\mathfrak{v}(A) = \mu$ and $\mathfrak{v}(B) = \top$. So $\mathfrak{v}(A \wedge B) \neq \bot$, i.e. the value of $A \wedge B$ is not false. The premise is not false here, but the conclusion is not true, for $\mathfrak{v}(A) = \mu$. Hence, $A \wedge B \nvDash A$, without changing the meaning of conjunction. We changed the logic (from logics validating those arguments to something else, for we changed the number of truth-values, the notion of validity and the separation of truth-values), but without clearly changing therefore the meaning of conjunction (the satisfiability condition was the usual one).

As with many ideas in philosophy, this last pluralist strategy surely has its shortcomings, but in order to discuss it in detail it is necessary to introduce further and more technical remarks on truth-values and the ways the collections of truth-values can be partitioned, as well as discussing whether we are left with something that is still logic after such changes. This will be done in Chapter 9.

Universal Logic

There is another argument, based on Jean-Yves Béziau's ideas on universal logic, for the idea that there is no logic holding in all cases. Béziau deploys a historical-practical-inductive-analogical argument to show that logical consequence doesn't need to satisfy any principles. Béziau calls this "the axiomatic emptiness of logic" (see Béziau, 1995, 2001, 2006, 2010). The argument can be reconstructed as follows:[3]

(B1) Virtually every theorem, principle for connectives, principle for the consequence relation, etc., let us call them collectively 'properties of a logic', has been thrown out or, at least, challenged.

[3] Although the same ideas recur here and there through the papers, the argument below has to be extracted by brute force, so to speak, from Béziau's writings, since it is never explicit and its parts appear as disconnected remarks, hidden amongst the discussion of several other topics. Just as an example, here are the views as expressed at some places by Béziau (2001):

"Traditionally the principle of contradiction is taken as a fundamental pillar of logic. The idea is that reasoning is not possible without it. Paraconsistency goes against this idea. And if paraconsistent logic is rightly a logic, therefore what are the ground principles of logic, if any?" (p. 5)

"The number of new logics has increased these last years due to the need of computer sciences, artificial intelligence, cognition, and all the stuff of our cybertime." (p. 20)

"My motivation and my terminology were taken from Birkhoff's famous notion of abstract algebra, that I found in *Lattice theory*, which is just a set with a family of operations. My idea was already that the basic foundations of logic were not more principles for the consequence relation than principles for connectives, like the principle of contradiction. I reached the idea that we must throw out all principles altogether, that logic *is not grounded on any principles or laws*." (p. 8; italics in the original)

(B2) The outcomes of such droppings and challenges have been regarded as logics.

(B3) If the properties $P_1, \ldots, P_n$ of a logic can be dropped or challenged, an additional property P_m also can be dropped or challenged and the result will still count as a logic.

(B4) The situation is analogous to the case of algebra, where an algebraic structure doesn't need to satisfy any property in particular.

(B5) Hence, a relation of logical consequence can be defined with no reference to a particular property of a logic.

Then Béziau proposes that a logic is a certain kind of mathematical structure and that a logical structure is a structure of the form $LS = \langle S, \vdash_{LS} \rangle$, where S is an arbitrary structure and $\vdash_{LS}$ is also an arbitrary relation on $\wp(S) \times S$ (provided we have means to obtain "powers" on such structure). Equivalently, it could be described as a pair $LS = \langle S, C_{LS} \rangle$, where C_{LS} is an arbitrary mapping $C_{LS} : \wp(S) \longrightarrow \wp(S)$. "Arbitrary" means here that neither the structure nor the relation or mapping need to satisfy any axiom. "Universal logic" would be a discipline whose subject matter is logical structures independently of the features of particular logics, analogous to universal algebra.

(B1) is a premise concerning the history of logic and, even if there were properties of a logic that have not been actually challenged, let us grant it. (B2) is a premise based on a kind of observation of practice. But practice might not be enough. It has been challenged, for example, by Quine (1970) and Dummett (1998), to name just two prominent philosophers, on the basis that certain logics are "uninterpreted theories" or "abstract algebras" rather than logics, receiving that name only by partial analogy. However, let us assume that such "partial analogy" suffices.

But even granting (B2), the problems of this argument are those of any inductive and analogical argument. On the one hand, one has to be very careful about how close the analogy between logic and algebra, as in (B4) and (B5), is. On the other hand, (B3) is ambiguous between a distributive and a collective sense of dropping properties and still obtaining logics. Thus, (B3) might mean at least the following two claims:

(B3′) If the properties $P_1, \ldots, P_n$ of a logic L_A can be dropped or challenged giving rise to another logic L_B, an additional property P_m can also be dropped or challenged from L_A and the result, L_C, will still count as a logic.

(B3″) If the properties $P_1, \ldots, P_n$ of a logic L_A can be dropped or challenged giving rise to another logic L_B, an additional property P_m can also be dropped or challenged from L_B and the result, L_C, will still count as a logic.

According to (B3′), for any property P of a logic, there could be another logic such that it lacks it. (B3″) says something stronger: there could be a logic such that it lacks any property of another logic, and even any properties of any other logic. And as we have seen from Beall and Restall's argument, if logical consequence is a kind of truth-preservation, (B3″) cannot be true, as reflexivity and transitivity would be required by the relation of truth-preservation. In order for (B3″) to be true, logical consequence should be a relation without specific requirements, i.e. axiomatically empty, but this was precisely what the argument was trying to establish. A different kind of argument would be needed then, and in Chapter 9 we will review some further attempts in the direction of axiomatic emptiness, closely connected to the remarks we made in the previous section.

Hartry Field's Pluralism about Epistemic Normativity

An altogether different approach to the issue of logical pluralism is Hartry Field's pluralism about epistemic normativity (Field, 2009a). As we will discuss in detail in the next chapter, logic is often considered to be normative for thought and reason. Logic tells us (perhaps via bridge principles: see the next chapter) which inferences we ought to draw and which consequences we ought to add to our beliefs. Field is an anti-realist about epistemic norms and doesn't believe that there is such a thing as a uniquely correct set of norms. At the same time, he takes the normative role of logic to be essential to logic (as opposed to truth-preservation). The goodness of a logic then depends on the quality of the epistemic system that it belongs to, and there is no reason to believe that there is a unique best goal-independent system.

Questions

1. Considering the different notions of pluralism (pointed, collected, distributed, universal) that we distinguished at the beginning of the

chapter: what does Shapiro's "eclectic pluralism" fall under? What about Field's epistemologically motivated pluralism?

2. Carnapian tolerance and Shapiro's eclectic pluralism have in common that all proposals (of languages or mathematical structures) should be looked at in light of their own logics. Can you think of drawbacks that such a tolerant pluralism might have?
3. For Field, logical pluralism is a consequence of his anti-realism concerning epistemic normativity. Assume instead realism about epistemic normativity. Is there a tension between such a position and (certain forms of) logical pluralism?

8 Logic, Reasoning and Rationality

At several occasions in this book, we referred to logic as a theory of the laws of thought and suggested that logic has normative implications for reasoning. For example, in Chapter 5, when discussing psychologism, we noted that one objection against a psychologistic approach to logic was that it couldn't account for the normativity of logic. If it is a psychological fact about Barbara that she reasons in a specific way, why should that have any implications about how she or anybody else *should* reason? This objection presupposed that logic is indeed in the business of telling us how we should reason. In this chapter we will look into what this actually means or could mean.

Why (Not) Reason Logically?

While there might be different views on how logic is normative for thought, and whether logic is itself to be considered a (partial) theory of reasoning, there is at least the following motivation for thinking that logic has *some* normative role to play for a theory of reasoning (see Steinberger, in press).

Let's assume that it is one of the aims of theoretical reasoning to arrive at an accurate representation of the world. In other words, we are interested in true beliefs about how things are. Beliefs have propositional content, content which will stand in logical relations. Since we have an interest in the truth of our beliefs, as at least a central part of our overall cognitive project, we should have an interest in these logical relations too, because they will clearly be of relevance. If I have some true beliefs, then their truth will carry over to their logical implications. If some belief of mine, on the other hand, implies a falsehood, then that belief can't be true. Finally, if some subset of my beliefs is inconsistent, then my representation of the world can't be accurate – at least one of my beliefs is false. These considerations motivate the following two principles:[1]

Logical Implication Principle (IMP) If S's beliefs logically imply P, then S ought to believe that P.

Logical Consistency Principle (CON) S ought to avoid having logically inconsistent beliefs.

(IMP) requires that I believe all consequences of my beliefs, whether or not they are consistent. (CON) requires that I don't believe inconsistencies, regardless of whether I believe the consequences of my beliefs. Hence, these are independent principles, both motivated from general, plausible considerations about the aims of cognition and theoretical reasoning. What could possibly be wrong with these?

As innocuous and plausible these might seem at a first glance, there are four central objections to these principles, put forward most forcefully by Gilbert Harman (1986).

The first objection observes that reasoning doesn't simply follow the pattern of logical consequence. You read on social media, by what appears to be a reliable news source, the headline that George W. Bush will receive the Nobel Peace Prize. You assume that *this is a reliable news source*, and you believe that *if this is a reliable news source, then George W. Bush will receive the Nobel Peace Prize*. Logic tells you that $P \rightarrow Q$ together with P entails Q. Should you now believe that George W. Bush will receive the Nobel Peace Prize? This conclusion seems so weird that it seems irrational to detach. You might rather want to revise P (perhaps you are in fact reading a cleverly disguised satirical magazine) or check whether it happens to be April Fool's Day and you shouldn't have believed the conditional.

In general, logic tells you what a set of statements entails. But just to go ahead and believe whatever turns out to be entailed by your other beliefs is not a good strategy. If you learn that your beliefs entail something that's strange or at odds with other things you believe, you should rather revise your other beliefs that entail the odd thing, than to simply add it to your "belief box". Florian Steinberger (in press) notes that the problem seems

[1] The formulation of these principles follows Steinberger (in press). Similar principles are discussed by Harman (1986).

to be that IMP requires us directly to believe the logical consequences. Perhaps you should update *somehow* when recognizing an entailment, but it seems to be wrong to require that the right way to update is in any case the addition of the entailment to the other beliefs. The instruction that IMP provides is too specific. Steinberger thus calls this objection the *Too Much Instruction Objection*.

There is another objection, which is very similar, due to John Broome (2000). As we know, every statement entails itself. Thus, if you happen to believe that today is a beautiful day, then, since this *logically implies* that today is a beautiful day, you also *ought* to believe that today is a beautiful day. But that seems wrong again. Not everything you believe is such that you ought to believe it (and it certainly doesn't become so just because you believe it).

The second objection also takes issues with IMP. IMP says that you ought to believe what your beliefs entail. But there's *lots* that your beliefs entail. Take the belief that you are shorter than 2 m. This entails that you are shorter than 2.01 m, and shorter than 2.02 m, 2.03 m, etc. It also entails that you are shorter than 2 m or Richard Montague was a founding member of *Black Sabbath*. There are endless trivial entailments of your beliefs, all of which would be entirely pointless for you to believe – why should you add them to your belief box and even bother to find out which they are? It seems that these additions would just clutter up your finite storage space and you would waste your cognitive resources deriving these entailments. Following Harman (1986) and Steinberger (in press), we call this the *Objection From Clutter Avoidance*.

A third objection elaborates further on requirements on cognitive resources that IMP and CON make. Since there are infinitely many consequences of our beliefs, finding out about all of them is just impossible. For just some of them it is also already impossible to find out whether they actually are entailments (perhaps the shortest proof of a certain claim from your other beliefs is so long that you wouldn't reach the end of it before the sun explodes). Likewise, there might be hidden inconsistencies in your belief set that are just impossible for you to ever detect. But, since 'ought' implies 'can', then CON and IMP must both be false.

Finally, Harman's fourth objection makes an observation that we have met already several times in this book. Sometimes there might be no way to avoid having an inconsistent set of beliefs. As already said above, inconsistencies might just be too hard to detect. Or perhaps you have detected that you have inconsistent beliefs, but you don't know what to do and/or don't have the resources available to find out how to restore consistency.

In the discussion of dialetheism in Chapter 10 we will encounter the idea that it might even be rationally required to have certain inconsistent beliefs. One way to see this (without assuming the existence of true contradictions) is to consider the *Preface Paradox*. Imagine you wrote a long book – perhaps your dissertation. All individual claims you make in the book are based on careful research. Let's say that the claims in the book are $P_1, \ldots, P_n$, and, for each P_i, you have good reason to believe it. Still, intellectual humility and plain common sense require that you also believe that, of all those many claims you made in the book, not all will be true. Human beings are fallible, and so are you. It is very unlikely that everything in the book is true. Let's call the claim that not all claims in your book are true Q. $\{P_1, \ldots, P_n\} \cup \{Q\}$ can't be consistent. However, it would be irrational (and pointless) now to drop some of the P_i. First of all, it's not clear how you are supposed to achieve that (given that you *do* believe each P_i), and secondly, unless you cut down on your beliefs to a very small set, the problem will just arise again. This speaks against CON, since it seems rationally permitted (if not required) to hold inconsistent beliefs in such a situation. (As Steinberger (in press) observes, it also speaks against IMP, because Q is transparently equivalent to the negation of the conjunction of your claims in the book, $\neg(P_1 \wedge \cdots \wedge P_n)$ – so you fail to believe an obvious logical consequence of your beliefs, their conjunction.)

Harman argues that IMP and CON fail to be plausible guidelines for updating your beliefs, because these principles rest on a fundamental confusion. IMP and CON confuse a *theory of implication and entailment* (i.e. a logic) with a *theory of reasoning*. The latter is, for Harman, supposed to be a theory of "reasoned change in view". Harman's paradigm example for the explananda of such a theory is the following:

> Intending to have Cheerios for breakfast, Mary goes to the cupboard. But she can't find any Cheerios. She decided that Elizabeth must have finished off the Cheerios the day before. So, she settles for Rice Krispies. In the process, Mary has modified her original intentions and beliefs. (Harman, 1986, 1)

On the basis of the information Mary received when finding the cupboard empty, Mary came to a reasoned change in view. What led to the change is a psychological process, and a theory of reasoning is supposed to be a normative theory that formulates principles or rules of revision that the relevant psychological processes need to follow in order to count as rational. What do such principles look like? Here are some of Harman's (1986, 55) examples:

Clutter Avoidance One should not clutter one's mind with trivialities.

Interest Condition (on Theoretical Reasoning) One is to add a new proposition *P* to one's beliefs only if one is interested in whether *P* is true (and it is otherwise reasonable to believe *P*).

Interest in the Environment One has reason to be interested in objects and events in one's immediate environment.

Interest in Fascilitating Theoretical Reasoning If one is interested in whether *P* is true and has reason to believe knowing whether *Q* is true would fascilitate knowing whether *P* is true, one has reason to be interested in whether *Q* is true.

Logic, of course, does not formulate such rules or principles about how one is supposed to revise one's beliefs. Logic formulates facts about implication. According to Harman, this observation holds also for proof-theoretic characterizations of logical consequence. After all, he observes, the introduction and elimination rules that, for example, the natural deduction[2] system defines, do not *prescribe* what one ought to do in a proof in order to reach the conclusion. In fact, when attempting a proof, we often begin our reasoning from the conclusion. Thus, the reasoning that goes into finding a proof is not itself properly described by the proof rules.

In an earlier paper, Harman (1984) flirts with, what he calls, the "extreme view" that logic is merely a body of truths, like any other descriptive science, and plays "no special role in telling what one may believe" (Harman, 1984, 109).[3]

[2] Although the inventors of the system, Stanisłav Jaśkowski (1906–1965) and Gerhard Gentzen (1909–1945), presumably did think that the rules follow the actual or natural reasoning of mathematicians.

[3] Harman's "extreme view" is an early version of *anti-exceptionalism*, the view that logic is just like any other science but with a more abstract subject matter. We will discuss this position in more detail in Chapter 10. Although Harman thinks that there are no serious objections to this view, he eventually does not endorse it in that paper.

Some of Harman's observations seem to be a bit beside the point, occasionally. For example, one should probably distinguish between the ways that mathematicians reason when constructing a proof and the ways in which we reason when constructing a proof in a natural deduction system. That the latter is perhaps not best described by the natural deduction rules themselves is perhaps not at issue. Also, there are systems that do formulate specific prescriptions what to do in each step of a proof (cf. proof procedures for tableau proofs in propositional logic). These systems have limited applications, but they exist for some systems.

However, it seems true that reasoning is a matter of adding new beliefs, but also of correcting beliefs. The latter is incompatible with the monotonicity of standard logic. Moreover, it is true that logic does not issue explicit prescriptions about what one ought to believe or how one ought to revise one's beliefs. Does that show that logic is irrelevant for reasoning?

Logical Constraints on Rationality

As Florian Steinberger (2017) notes, there are two ways to answer Harman's challenge. One way could be to criticize either Harman's conception of logic or his conception of reasoning for being too narrow. For example, one could point out that there are several theories of belief revision that take logical considerations as their starting point or indeed are logics (non-monotonic logics, dynamic doxastic logics). Some of these theories do talk about mental states and say what one ought to believe. Moreover, they typically give up the monotonicity of standard logic and allow for the actual revision of beliefs.

Likewise one could find fault with Harman's conception of reasoning. Harman considers the issue from an internalist, first-person perspective. From that perspective, logic doesn't seem to provide proper guidance for belief revision. But perhaps the better perspective is externalist. From that perspective it might well be that belief transitions need to follow logical principles in order to count as rational.

A second strategy for answering the challenge takes Harman's conception of reasoning and logic for granted, but tries to show that the latter is relevant for the former, by formulating *bridge principles*. How do bridge principles work? Let's consider an example from ethics (which we have already applied at the beginning of this chapter). As is usually accepted, it would be a mistake (actually, an instance of the *naturalistic fallacy*) to derive normative claims straightforwardly from descriptive claims. That doesn't mean, however, that the way the world actually is is irrelevant for how it ought to be. For example, it seems to matter, for whether a moral obligation is legitimate, that the addressee of the obligation is able to follow it. Moral obligations shouldn't require superhuman tasks, for instance. But what counts as "superhuman" is, of course, a descriptive question. Thus, the bridge principle *ought implies can* bridges the gap between descriptive facts (for example, facts about what humans are normally capable of), which, by themselves, have no normative force or normative implications, and the validity of normative requirements.

Such bridge principles for logic have recently been discussed in some detail, starting with an unpublished paper by John MacFarlane (2004), followed by papers by Hartry Field (2009a), Florian Steinberger (in press) and others. MacFarlane brought new life to the discussion of the normative role of logic by, first of all, developing a template that allows us to map the logical space for the formulation of such bridge principles. The general template for MacFarlane is the Bridge Principle (MacFarlane, 2004):

Bridge Principle If $A, B \vDash C$, then (normative claim about believing A, B, and C)[4]

The *consequent* of instances of this template can now vary with respect to three parameters:

1. The *type of the deontic operator* may vary between *obligation*, *permission* and (defeasible) *reason for belief*.
2. The *polarity* of the obligation/permission/reason for belief may vary between a positive polarity (reason to believe) and a negative polarity (reason not to disbelieve).
3. The *scope* of the operator may also vary between applying only to C, or the conditional $A \wedge B \to C$ as a whole or distribute over that conditional.[5]

This leads to the the options shown in Table 8.1.

[4] In this template, 'A', 'B' are shorthand for any number of premises.
[5] To be exact, the operator would not apply to that conditional, but some conditional invoking the agent's doxastic attitudes towards those propositions, e.g. 'You ought to (believe C, if you believe A and B)'.

Table 8.1 *MacFarlane's bridge principles (from MacFarlane, 2004)*

If $A, B \vDash C$, then ...

- C Deontic operator embedded in consequent.
  - o Deontic operator is strict obligation (ought).
    - Co+ ... if you believe *A* and you believe *B*, you ought to believe *C*.
    - Co– ... if you believe *A* and you believe *B*, you ought not to disbelieve *C*.
  - p Deontic operator is permission (may).
    - Cp+ ... if you believe *A* and you believe *B*, you may believe *C*.
    - Cp– ... if you believe *A* and you believe *B*, you are permitted not to disbelieve *C*.
  - r Cr+ ... if you believe *A* and you believe *B*, you have reason to believe *C*.
    - Cr– ... if you believe *A* and you believe *B*, you have reason not to disbelieve *C*.
- B Deontic operator embedded in both antecedent and consequent.
  - o Deontic operator is strict obligation (ought).
    - Bo+ ... if you ought to believe *A* and you believe *B*, you ought to believe *C*.
    - Bo– ... if you ought to believe *A* and you believe *B*, you ought not to disbelieve *C*.
  - p Deontic operator is permission (may).
    - Bp+ ... if you may believe *A* and you believe *B*, you may believe *C*.
    - Bp– ... if you may believe *A* and you believe *B*, you are permitted not to disbelieve *C*.
  - r Br+ ... if you have reason to believe *A* and you believe *B*, you have reason to believe *C*.
    - Br– ... if you have reason to believe *A* and you believe *B*, you have reason not to disbelieve *C*.
- W Deontic operator scopes over whole conditional.
  - o Deontic operator is strict obligation (ought).
    - Wo+ ... you ought to see to it that if you believe *A* and you believe *B*, you believe *C*.
    - Wo– ... you ought to see to it that if you believe *A* and you believe *B*, you do not disbelieve *C*.
  - p Deontic operator is permission (may).
    - Wp+ ... you may see to it that if you believe *A* and you believe *B*, you believe *C*.
    - Wp– ... you may see to it that if you believe *A* and you believe *B*, you do not disbelieve *C*.
  - r Wr+ ... you have reason to see to it that if you believe *A* and you believe *B*, you believe *C*.
    - Wr– ... you have reason to see to it that if you believe *A* and you believe *B*, you do not disbelieve *C*.

There are some comments we should make about this table. First of all, the table gives us a handy way to refer to each of the 18 options by their short name. In the short name, the first letter indicates the scope of the operator, the second indicates the type of the operator and the third symbol the polarity.

The table does not represent the fact that agents may or may not know that the antecedent of the conditional is the case, i.e. that $A, B \vDash C$. One could represent this by adding a further letter, 'k', for versions of the bridge principles in which the agent *knows* that $A, B \vDash C$. Thus, Wr–k would be

Wr–k If you *know* that $A, B \vDash C$ you have reason to see to it that if you believe A and you believe B, you do not disbelieve C.

Adding this option to the table would generate 36 different bridge principles. Steinberger (2017) notes that other, non-factive attitudes might also be plausible candidates for modifications of the antecedent of the conditional. Consider an agent who strongly believes that $A, B \vDash C$, while in fact there is no such implication. It seems that such an agent would be even more irrational if she failed to revise her beliefs in accordance with the strongly believed entailment.

Other complications could be added by using different deontic operators in the cases under the Bs. Now the operator is simply distributed over the conditional, but one might consider a different operator for antecendent and consequent. For example, a variant of Bo+ could be '. . . if you *ought* to believe A and you believe B, you *have reason* to believe C' (cf. Steinberger, 2017).[6]

MacFarlane (2004) suggests some considerations one might apply when evaluating these options. Some of these follow from the different objections that Harman and others had already formulated against simplistic principles, such as IMP or CON, and indeed IMP is represented here as Co+. Remember, for example, Broome's objection against IMP: every belief implies itself, but simply because you happen to believe A it doesn't follow that you ought to believe A. It also doesn't follow that you ought not disbelieve it or may believe or are permitted not to disbelieve it. Thus, these considerations eliminate Co+, Co–, Cp+, Cp– and also their -k variants.

[6] As Steinberger also notes, the plausible versions of the mixed cases will have a deontic operator in the antecendent that is at least as strong as the operator in the consequent.

Other considerations that MacFarlane discusses are *Excessive Demands* (belief in some axioms, for example, shouldn't require that you therefore believe in all theorems that follow from them), and that bridge principles should give some plausible prescription for the Preface Paradox (for example, they shouldn't require that you ought to avoid inconsistency at all costs, but perhaps merely recommend revising your beliefs when encountering an inconsistency). Further, such principles should pass the *Strictness Test* (i.e. the relation between believing something and believing its consequence is strict: one can't believe the former but not the latter and be as one ought to be) and avoid *Logical Obtuseness*: logical implications need to have some impact on our beliefs, we can't simply stay neutral with respect to *B*, if *B* follows from *A*, and we believe *A and* be as we ought to be. Likewise, we can't *improve* our epistemic situation by just being more ignorant about the logical facts (MacFarlane calls this the *Priority Question*); the -k variants that were supposed to liberate us from excessive demands conditionalize on our logical knowledge, which has as a consequence that if we know less about logical implication, we have more freedom to believe inconsistent crap.

> But this looks backwards. We seek logical knowledge so that we will know how we ought to revise our beliefs: not just how we *will* be obligated to revise them when we acquire this logical knowledge, but how we are obligated to revise them even now, in our state of ignorance. (MacFarlane, 2004)

MacFarlane himself ends up endorsing Wo– and Wr+ (and the principles they imply, such as Wo–k, Wr+k, Wr– and Wr–k). But he notes that the considerations that he suggested for evaluation do not always pull in the same direction.

Problems with Bridge Principles

Perhaps you noticed this already when going through the considerations above. The Strictness Test on the one hand, and Excessive Demands and Preface Paradox on the other hand, pull in opposite directions; the Strictness Test seems to support ought-based principles, while the other two seem to caution us against such strict principles.

As Florian Steinberger (in press) argues, this somewhat unfortunate situation can be improved when we distinguish between the several normative roles that a principle may play. In order to distinguish these roles we need to think about why we want to have a principle. What is its function supposed to be?

One function that we have implicitly assumed in this chapter is *guidance*; we want to know from the first-person point of view how we ought to revise our beliefs. For example, we criticized IMP for being a bad *directive* in this sense. It would lead us to clutter our minds if we were to follow its prescription.

Another function is *evaluation*: we want to use the principle to learn whether a certain belief revision was good or bad or correct or incorrect. Finally, we might use principles for *appraisal*. Is an agent to blame or to praise for a specific belief revision or a failure to revise her beliefs?

These three functional roles, *directions*, *evaluations* and *appraisals*, are, of course, not independent. Directions should not recommend what isn't good from an evaluative point of view, and following directions should make one a candidate for positive appraisal. However, the roles are also clearly distinct. For appraisal it should matter whether an agent recognized a logical implication, while this might not matter for an evaluation of the belief revision itself. By distinguishing these different roles, Steinberger distinguishes different bridge principles (in addition to the distinctions introduced above, one further parameter for each of the three functional roles). However, now we need to evaluate the principles according to their roles. Thus, the consideration we went through above would be weighed differently, depending on which role the principle has that the considerations are applied to, and stop pulling in different directions.

Mono-Agent and Multi-Agent Logic

In a recent paper, Catarina Dutilh Novaes (2015) suggests another normative function that bridge principles might play and thus a further normative role for logic. She agrees with Harman that logic is not normative for individual belief revisions. But she observes that, traditionally, logic was developed in the context of dialogical arguments, special, relatively regulated forms of debates. These dialogical origins of logic began to fall into oblivion with Descartes.

While in the original tradition debates where seen as an interplay between a proponent (establishing premises and defending a claim) and an opponent (who tries to prevent the proponent from making further steps), modern logic became a mono-agent affair that internalized the opponent's perspective. However, the normative force of logic, according to Dutilh Novaes, can only be properly appreciated in the original dialogical perspective. For that perspective she formulates two new, multi-agent bridge principles (as continuations of 'If $A, B \vDash C$, then ...'):

Wo+d_o ... Opponent ought to see to it that, if he has granted A and B and Proponent puts forward C, then he will grant C.

Cp+d_p ... if Opponent has granted A and B, then Proponent may put forward C (and require Opponent to grant it).

Here the subscripts 'o' and 'p' indicate the addressee of the normative recommendation. Indeed this reinterpretation of logic circumvents the objections by Harman; however, it does so at the cost of making logic irrelevant for (individual) rationality and reasoning.

Logic, Knowledge and Information

In the context of discussing logic and rationality, we should also address an issue that relates to our knowledge of logical truths and our grasp of logical relations among our beliefs. In the chapter on epistemology (Chapter 6) we have already discussed whether we should think of logic as being *a priori* knowable. In the current chapter we took it for granted that ordinary people do not just know all the logical truths and do not simply already know what their beliefs entail. If the latter weren't the case, there wouldn't be much point in discussing the normativity of logic. As you remember, Gilbert Harman's objection was that IMP, i.e.

Logical Implication Principle (IMP) If S's beliefs logically imply P, then S ought to believe that P.

formulates an unreasonable norm. We would clutter our mind with useless beliefs, if for every proposition P that we believe, we also had to believe $P \vee Q_1$, $(P \vee Q_1) \vee Q_2$, $((P \vee Q_1) \vee Q_2) \vee Q_3$, etc. This objection presupposes that we don't automatically know all the implications of P when we believe P; we have to engage in *deductive reasoning* in order to learn what these implications are. This is why IMP could possibly lead to a waste of time and resources. On the flip side, if we always already knew the logical implications of all of our beliefs, IMP would not burden us with any intellectual extra work or lead to a cluttering of our minds, but would just be an empty norm: it would be impossible not to comply with it.

So it seems that – for the discussion of the normativity of logic to make any sense in the first place – we need an account of *a priori* knowability and knowability via deductive reasoning in particular that is compatible with the idea that such reasoning leads to *new* knowledge; that we *learn* something when we engage in deductive reasoning, that we acquire new beliefs.

As with so many topics in this book, this issue is also connected to several other deep questions in philosophy, most of which we can't discuss here because of time and space constraints (let alone our incompetence). For example, the question of how we should deal with the question of how we could possibly acquire new beliefs via deductive reasoning is closely related to the question of whether and how we can develop an epistemic logic (a logic of knowledge and belief), as well as with the question of how we should best think of the contents of propositional attitudes.

In epistemic logic one could begin by introducing a logical constant 'K' that behaves syntactically and (to some extent) semantically like the '$\Box$' of modal logic. Thus, $\ulcorner$it is known that $\phi\urcorner$ (perhaps by some fixed agent *S*) could be formalized as $\ulcorner K\phi\urcorner$. Treating the knowledge operator like a necessity operator has certain desirable implications. However, if we take the logic of 'K' to be a "normal"[7] modal logic, then $\ulcorner K\phi\urcorner$ implies $\ulcorner K\psi\urcorner$ whenever ϕ implies ψ. But this seems wrong for the reasons provided above: people don't automatically know whatever is logically entailed by their knowledge.

Contents of propositional attitudes and propositions in general are likewise fruitfully thought of as being sets of possible worlds. On such an account the content of what Jim believes when he believes that Jill is at

[7] A *normal* modal logic can syntactically be characterized as a set of formulas that contains all tautologies of propositional logic, all instances of the axiom schema

$$\mathbf{K}\colon \Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$$

and is closed under *modus ponens* and the *rule of necessitation*, according to which $\vdash \phi$ implies $\vdash \Box\phi$.

(The name of the axiom schema is unrelated to our operator K but derives from 'Kripke', because Saul Kripke developed a very elegant way to study such normal modal logics semantically. See Sider (2010) or Hughes and Cresswell (1996) for more details.)

home is modelled as the set of all possible worlds in which Jill is at home. But in all those possible worlds in that set, all of the logical implications of Jim's belief will also be the case (as, for example, that Jill is at home or Bob is in Barcelona; that Jill is at home, or Bob is in Barcelona, or Carla is in Vancouver, etc.). Thus modelling the contents of propositional attitudes as sets of possible worlds will again lead to *logical omniscience*: we should know all logical implications of our beliefs (which includes, of course, all logical truths themselves).

One could see this as a technical problem for which the respective model is to blame. Perhaps epistemic logics should be better modelled with non-normal modal logics, and perhaps the contents of propositions can't be identified with something as coarse-grained as a set of possible worlds. However, in order to know whether and how to adjust the model, we need to answer a philosophical question first: what exactly do we learn when we engage in deductive reasoning?

Michael Dummett expressed the philosophical puzzle underlying this question brilliantly in his essay "The justification of deduction":

> The existence of deductive inference is problematic because of the tension between what seems necessary to account for its legitimacy and what seems necessary to account for its usefulness. For it to be legitimate, the process of recognising the premises as true must already have accomplished whatever is needed for the recognition of the truth of the conclusion; for it to be useful, a recognition of its truth need not actually have been accorded to the conclusion when it was accorded to the premises. (Dummett, 1978, 297)

The idea of the usefulness of a deductive proof concerns the fact that we engage in deductive reasoning for a purpose. We want to learn something new, which we didn't know before, namely *that the conclusion is true*. The legitimacy of a deductive inference can be understood in two ways. It can be understood as a matter of *subjective entitlement*. When do we know that we are allowed to draw an inference? If it was required for legitimately drawing an inference that we first *recognize* that the truth of the conclusion is entailed in the truth of the premises, then this seems to be in direct tension with the purported usefulness of such an inference. We would already need to know what we want to find out by the inference as a precondition for making the inference. But we can just take this observation as an argument against this requirement. For a deductive inference to be legitimate it can't be required that we recognize beforehand that the truth of the conclusion is entailed in the premises.

But even if the legitimacy of a deductive inference is not understood in that subjective sense, but in the sense that in order for a deductive inference to be legitimate, the truth of the conclusion must be entailed in the truth of the premises (whether we recognize this or not), there is still a tension with the inference's usefulness. The best way to see this is to think about it in terms of the information we gain when we draw a deductive inference.

On a standard analysis, we can explicate the information gained by a piece of new knowledge as the *doxastic possibilities* that the epistemic agent can eliminate. Doxastic possibilities are the ways the world might be given what you believe. Let's assume that you are searching for your office keys. For all you know, they might be at your parents' place (who you visited over the weekend) or they might be somewhere in the house, either on your desk or in the kitchen. These are three different ways the world might be for all you know that differ with respect to the location of your keys. You know something about where your keys are, because you can eliminate some possibilities concerning their whereabouts. (For all you know, they can't be on the moon.) But you don't know where they are exactly.

Now, your roommate tells you that she saw the keys somewhere in the house today. That allows you to eliminate a further possibility, namely that the keys are at your parents' place. Your friend's remark is informative because it allows you to eliminate that possibility. According to the standard view, there is also an *inverse relationship* between the increase of available information and the decrease in doxastic possibilities (cf. Barwise, 1997, 491). For example, assume your roommate informed you instead that she was just in the kitchen and the keys weren't there but that she saw them somewhere in the house today. That excludes the possibility that they are at your parents' place, but also that they are in the kitchen. Since this piece of information allows you to eliminate more doxastic possibilities, it is more informative than to learn merely that the keys are somewhere in the house. In fact, now there are *no* alternative possibilities for your keys to be other than that they are on your desk. Knowing where your keys are entails that the doxastic possibilities for their whereabouts are reduced to one.

Let us apply that analysis to deductive reasoning. How does information increase when we reason deductively from premises to conclusion? In particular, what is the informational content associated with learning that the conclusion is true?[8]

As we know, a deductive inference is valid (and thus legitimate) if there are no possibilities that make all the premises true, but fail to make the conclusion true. In other words, the set of possibilities there are for all you know when you know the premises, and the set of all possibilities there are for all you know when you know the premises *and* the conclusion is the same. There is no difference between these two sets of possibilities associated with the two information states: knowing the premises vs. knowing the premises and the conclusion. Hence, you don't learn anything from a deductive inference if the inference is legitimate. But then deductive reasoning is plainly useless.

It is easy to see that, under the same assumptions, we should also all come out logically omniscient. Since the set of possibilities that logical truths exclude is empty (a logical truth is true in all possibilities), we don't need to rule out any possibility in order to know a logical truth. We immediately know all of them. What should we say about this?

The first option could be to simply deny that we ever really learn anything new from deductive inferences. On that view, in most cases in which we seem to make a deductive inference we indeed make a deductive inference, but nothing is in fact learned on these occasions. How could that be? Perhaps our psychology just plays a trick on us. Perhaps a deductive inference only brings to our attention what wasn't yet in focus (albeit known).

Some views that consider logical omniscience an idealization of a theory of rationality fall into this first category. Perhaps deductive reasoning is not a matter of finding out something new, but a matter of retrieving information we already possess. However, since we are finite, slow beings,

8 This question assumes that what we learn from a deductive inference is that the conclusion is true in addition to already knowing that the premises are. An alternative account could consider the content of what we learn from such an inference to be that the premises are incompatible with the negation of the conclusion. On the standard account this makes no difference. Cf. Jago (2013).

this information retrieval takes time and sometimes leads to mistakes. We then mistake this process for the process of learning something new.

Something like this is what Carnap and Bar-Hillel (1952) had in mind when developing their theory of semantic information. They allow for a notion of "psychological information" to be assigned, for example, to propositions of mathematics, but hold on to the view that the informational content of such propositions for an "ideal receiver" would nevertheless be zero. But what would be the motivation for making such an idealization? Is it because we believe that actual agents would be logically omniscient if only they were undisturbed by other factors? What is the evidence for such an assumption? Is the idealization a harmless simplification in an otherwise approximately true theory of epistemic logic? As Stalnaker (1991) argues, this idealization seem ill-motivated in light of the fact that it renders all information processing or computation wholly unintelligible, although these seem to be activities that are essential to rationality and cognition.

Maybe the proper interpretation of the idealization is the purported normativity of the theory of rationality that a logic of knowledge and belief should contribute to. However, as we have seen above already, to believe all the logical consequences of what you belief is to clutter your mind with useless content. Stalnaker (1991) argues that the best interpretation of such an idealization would be that we simply don't know how to develop a theory of knowledge that wouldn't have the consequence that epistemic agents are logically omniscient. This, of course, would be a rather pessimistic motivation. Thus it's worth looking for alternative solutions to the problem of deduction.

A second option would resolve the tension between the usefulness and the legitimacy of deductive inferences by analysing those cases in which we seem to gain new knowledge as cases in which we are not actually making a proper deductive inference. For all proper deductive inferences it is true that we can't learn anything new, but those inferences where we learn something new merely *appear* to be deductive. What could cause the misleading appearance? Well, perhaps in the process of the inference the meaning of the premises gets changed, such that the premises eventually deductively entail the conclusion although they initially didn't. In retrospect we might be unable to detect this, since we are now attaching the new meaning to the premises.

Michael Dummett (1978) ascribes such a view to Ludwig Wittgenstein, who purportedly endorsed it in his *Remarks on the Foundations of Mathematics* (Wittgenstein, 1983):

> In that book he held that a proof induces us to accept a new criterion for the truth of the conclusion. [...] When the proof is given that a cylinder intersects a plane in an ellipse, we acquire a new criterion for a plane's figure being an ellipse [...] We have modified the meaning of the statement of the theorem, so that, in our example, the adoption of the new criterion for its application modifies the meaning that we attach to the predicate 'ellipse'. [...] [Wittgenstein's thesis] must be understood as involving that there are, or may be, plane figures formed by the intersection of a cylinder with a plane which could not have been recognised as ellipses before the proof was given. (Dummett, 1978, 300–301)

By adding a theorem about ellipses to our knowledge, we change the meaning of 'ellipse'. But then the conclusion of the proof was obviously not "contained" in the meaning of the premises. The premises, by themselves, would have allowed that there are mathematical objects that we can now, after the proof, recognize as ellipses, that weren't ellipses before.[9] As Dummett points out, this makes room for the idea that we learn something new through a proof, but at the cost of losing an account of a proof's legitimacy and persuasiveness.

Dummett contrasts this view with a "modified" Wittgensteinian view which agrees that inferential relations between statements constitute the meaning of the expressions within the statements (such that the inferential relations of statements with the term 'ellipse' with other statements constitute the meaning of that term), but denies that new inferential connections are *added* when a new proof is made. The connections were there all along in the language as a whole. Deduction on this conception is then "useful" because it allows us to reach conclusions in this interconneced web of statements we wouldn't have been able to reach without it. But this

[9] Wittgenstein's view is perhaps better seen as a description of the progress of mathematics. As Imre Lakatos (1976) shows, mathematicians have indeed some room to decide how they want to fix the intended interpretation of certain expressions by deciding on what counts as a proof and what counts as a counterexample to a theorem. But, Lakatos' point would then best be put as showing that mathematics, despite a common conception of it, does not in fact just proceed by deductive proof.

view still fails as an account of what *information* we gain by a deductive inference, if the conclusion doesn't rule out possibilities that weren't already ruled out by the premises.

On the third option, one would modify the information measure to help with exactly that problem. We just assumed when setting up the puzzle of deduction that the possibilities that we talk about in the definition of entailment (that there are no possibilities that make the premises true but fail to make the conclusion true) are the same possibilities as those doxastic possibilities that we talk about in the analysis of information. Accounts following this strategy could claim that there are more doxastic possibilities than logical possibilities, and that those doxastic possibilities get eliminated in (certain cases of) deductive reasoning and should be taken into account when measuring the information increase. (A variant of the third approach would define a second kind of content that is irrelevant for entailment, but is a relevant ingredient for the analysis of informational content.) The third strategy can be fleshed out in various ways, depending on how one models the doxastic possibilities.[10] These accounts are quite technical, and explaining them in any detail would go beyond the scope of this book, but we can explain one such account in the abstract to show what would be at issue for this third strategy to work.

We will take as an example a proposal by Jaakko Hintikka (1970).[11] Hintikka observed that we can distinguish formulas of first-order predicate logic by their "quantificational depth". One can think of quantificational depth as a function of the interaction between universal and existential quantifiers, once a formula is brought into some canonical form that would allow us to compare formulas in this way. In derivations, computational complexity can increase. To simplify, let's assume that when it does, we consider the proof to be information increasing (and we consider logical truths of the form $\ulcorner\phi \rightarrow \psi\urcorner$ to be informative if ψ has a greater quantificational depth than ϕ).

Such a theory would allow us to make some distinction between non-informative proofs and logical truths and informative ones. It would also deliver some measure of informativeness. However, it is not clear how one

10 A good overview of strategies of that kind can be found in Jago (2013).

11 For a critical discussion, see Bremer and Cohnitz (2004) and Sequoiah-Grayson (2008).

should go about validating such a theory. How should we check whether the measure proposed does indeed track (some sort of) informativeness? (Are we supposed to have intuitions about this? Should we validate it with the help of empirical psychology?) This relates closely to a second problem with the account: this measure is only applicable to polyadic predicate logic (since, for monadic predicate logic, or propositional logic, there are no differences in quantificational depth). Should we conclude from this that inferences at those levels are always uninformative? Intuitively that doesn't seem right.

In defence of Hintikka's account one could argue that the apparent information increase that takes place when reasoning in a way that is representable in propositional logic (or monadic predicate logic), is only a psychological phenomenon that doesn't represent any objective increase in information. But this move would just get us back to the problem of justifying this idealization (and the same considerations we went through above would apply again).

These problems will presumably apply to any formal account that tries to measure the objective amount of information gained in a valid inference on the basis of formal properties of premises and conclusion.[12]

Questions

1. We suggest that Harman's criticism of the idea that logic can play a normative role can be (partly) deflected if we take an externalist perspective and *evaluate* the belief-forming of an individual. To what extent can such a move deal with Harman's objections?
2. Consider Dutilh Novaes suggestion for understanding the normativity of logic as applied to multi-agent dialogues. Can you think of objections to Wo+d_o or Cp+d_p?
3. In the second part of the chapter we looked at ways to make sense of the idea that we gain new knowledge through deductive reasoning. Perhaps this is just completely misguided. Which consequences of a belief

[12] For a discussion of further alternative approaches to the problem, such as ontological solutions, or solutions that make use of measures of computational complexity, see Bremer (2003).

(or a set of beliefs) an agent draws seems pretty much a straightforward question of empirical psychology. Why should we be able to find a formal answer to this question? But then the information that an agent gains through deductive reasoning seems equally a matter of empirical psychology, right?

9 Beyond Truth-Preservation

As we have seen, logical consequence or logical validity has usually been characterized as necessary truth-preservation from premises to conclusion(s), or necessary *forwards truth-preservation*: in any case in which the premise(s) is (are) true, the conclusion(s) is (are) true as well, which implies that logical consequence is at least *reflexive* and *transitive*. Recently there have been a number of considerations that seem to go against these entrenched ideas. These considerations can be grouped into four main classes:

- First, there are research programmes seeking the best solutions to self-reference paradoxes, and all the best prospects involve rejecting some of the usual Tarskian properties of logical consequence, reflexivity or transitivity, but then forwards truth-preservation cannot be what characterizes logical consequence.
- Second, and also paradox-related, is the idea that the very notion of logical validity itself is plagued by paradoxes similar to the more well-known paradoxes that afflict notions such as *truth* or *set*. Hence, logical validity cannot be characterized as (forwards) truth-preservation.
- Third, there is a project connected to inferentialism that uses concepts corresponding to cognitive states like *acceptance* or *rejection* – or their linguistic correlates, like *assertion* or *denial* – instead of truth or falsity, in order to define logical validity. But then, what is forwards-preserved is something other than truth, and usually the resultant preservation relations are neither reflexive nor transitive.
- Fourth, there are attempts to generalize the usual notions, for a wide range of reasons, but we will concentrate here on those that attempt to identify *universal* features of logical consequence.

Each of the following sections will focus on one of these groups of ideas that challenge logical consequence as characterized by truth-preservation.

Solutions to the Paradoxes of Self-reference

According to tradition, logical consequence is a kind of truth-preservation, and truth-preservation is a relation that is at least reflexive and transitive; thus, a notion of logical consequence that is either non-reflexive or non-transitive can seriously be doubted to be a relation of truth-preservation. Nevertheless, many logicians now think that in order to give a uniform solution to the paradoxes of self-reference one must revise logic by dropping one of the usual Tarskian properties of logical consequence, such as reflexivity or transitivity. Most of these approaches drop the properties via their sequent-calculi expressions: structural rules.

We will call 'non-Tarskianists' those willing to give up a Tarskian property of logical consequence in order to solve paradoxes of self-reference (cf. Restall, 1994; Petersen, 2000; Weir, 2005; Zardini, 2011; Ripley, 2013; French, 2016); those who would instead prefer modifying the conception of truth, set, any connective or, in general, something other than the Tarskian properties of logical consequence in order to deal with paradoxes will be dubbed 'Tarskianists' (cf. Priest, 2006; Field, 2008; Beall, 2009; Halbach, 2011; Scharp, 2013). We will show here that non-Tarskianists might very well have a good case for their view, and logical consequence might well be something other than a species of truth-preservation.

Suppose, as allowed by diagonal mechanisms for self-reference, that we have a sentence in our language, call it 'λ', that is identical to $\neg T\langle\lambda\rangle$, where $\langle\lambda\rangle$ is a name for the sentence λ and T is a truth predicate. Let us also consider the following, very plausible, rules governing T:

$$\text{TR}\ \frac{\Gamma \Rightarrow \phi, \Delta}{\Gamma \Rightarrow T\langle\phi\rangle, \Delta} \qquad \frac{\Gamma, \phi \Rightarrow \Delta}{\Gamma, T\langle\phi\rangle \Rightarrow \Delta}\ \text{TL}$$

These rules seem to capture our intuitive understanding of the behaviour of 'is true', allowing us to derive the Tarski-biconditionals (that $T\langle\phi\rangle$ if and only if ϕ) in the presence of the usual rules for the conditional. This means that truth is "transparent", as one can then intersubstitute $T\langle\phi\rangle$ and ϕ anywhere in a proof, for any ϕ. But this is more than enough to get us into trouble, using only a small number of the rules we gave in Chapter 2:

$$
\begin{prooftree}
\AxiomC{$\lambda \Rightarrow \lambda$}
\LeftLabel{substitution}
\UnaryInfC{$\neg T\langle\lambda\rangle \Rightarrow \lambda$}
\LeftLabel{$\neg$ L}
\UnaryInfC{$\Rightarrow T\langle\lambda\rangle, \lambda$}
\LeftLabel{TR}
\UnaryInfC{$\Rightarrow \lambda, \lambda$}
\LeftLabel{CR}
\UnaryInfC{$\Rightarrow \lambda$}
\AxiomC{$\lambda \Rightarrow \lambda$}
\RightLabel{substitution}
\UnaryInfC{$\lambda \Rightarrow \neg T\langle\lambda\rangle$}
\RightLabel{$\neg$ R}
\UnaryInfC{$\lambda, T\langle\lambda\rangle \Rightarrow$}
\RightLabel{TL}
\UnaryInfC{$\lambda, \lambda \Rightarrow$}
\RightLabel{CL}
\UnaryInfC{$\lambda \Rightarrow$}
\RightLabel{Cut}
\BinaryInfC{$\Rightarrow$}
\RightLabel{WL}
\UnaryInfC{$\phi \Rightarrow$}
\RightLabel{WR}
\UnaryInfC{$\phi \Rightarrow \psi$}
\end{prooftree}
$$

To be clear: the problem here is that ϕ and ψ are completely arbitrary. So the conclusion we've arrived at is that anything at all follows from anything else at all. And that's absurd. Let's call this the *Liar Paradox*.

Here is a different problem, which is called the *Curry Paradox*: Let ϕ and ψ be any sentences in our language whatsoever and suppose, again as allowed by diagonal mechanisms for self-reference, that we have a sentence in our language, call it 'C', that is identical to $C \rightarrow (\phi \rightarrow \psi)$. If we suppose, as is entirely plausible, that the conditional obeys the following two-way rule[1]

$$
\begin{array}{c}
\Gamma \Rightarrow \phi \rightarrow \psi \\
\hline\hline
\Gamma, \phi \Rightarrow \psi
\end{array}
\; \rightarrow \text{I/E}
$$

then, if we have Contraction, it turns out that we are in just as much trouble as we were with the Liar:

$$
\begin{prooftree}
\AxiomC{$C \Rightarrow C$}
\LeftLabel{substitution}
\UnaryInfC{$C \Rightarrow C \rightarrow (\phi \rightarrow \psi)$}
\LeftLabel{$\rightarrow$ E}
\UnaryInfC{$C, C \Rightarrow \phi \rightarrow \psi$}
\LeftLabel{CL}
\UnaryInfC{$C \Rightarrow \phi \rightarrow \psi$}
\LeftLabel{$\rightarrow$ I}
\UnaryInfC{$\Rightarrow C \rightarrow (\phi \rightarrow \psi)$}
\LeftLabel{substitution}
\UnaryInfC{$\Rightarrow C$}
\AxiomC{$C \Rightarrow C$}
\RightLabel{substitution}
\UnaryInfC{$C \Rightarrow C \rightarrow (\phi \rightarrow \psi)$}
\RightLabel{$\rightarrow$ E}
\UnaryInfC{$C, C \Rightarrow \phi \rightarrow \psi$}
\RightLabel{CL}
\UnaryInfC{$C \Rightarrow \phi \rightarrow \psi$}
\RightLabel{Cut}
\BinaryInfC{$\Rightarrow \phi \rightarrow \psi$}
\RightLabel{$\rightarrow$ E}
\UnaryInfC{$\phi \Rightarrow \psi$}
\end{prooftree}
$$

The first main advantage of non-Tarskian approaches is their power to maintain, without restriction, the operational rules governing each piece of familiar logical vocabulary. As we have already seen, in the presence of Reflexivity, Contraction, Cut and transparent truth, the negation rules suffice to cause problems via the Liar Paradox. The conditional rules do as well – this time with even fewer assumptions required.

[1] It's not strictly necessary that we add this rule to our system – the problem we point out below arises in the system introduced in Chapter 2. But it arises in a proof that fits better on the page if we adopt the rule, so we do so.

Things are not so simple for Tarskian approaches to transparent truth. Such an approach must handle the Liar Paradox by modifying the negation rules somehow; that is the only remaining degree of freedom. But such a modification has usually nothing to say about the Curry Paradox, where negation is not involved. Similarly, it must handle the Curry Paradox by modifying the conditional rules; but this says nothing about the Liar.

This leads to two problems. First, it opens the Tarskianists to a familiar charge, namely that they are not really addressing the paradoxes at all, but only changing the subject. After all, so the objection has it, the Liar Paradox we originally cared about is the one formulated with a connective '¬' obeying the usual negation rules. A connective that does not obey the negation rules simply is not that same connective, no matter what symbol we use to write it (cf. our discussion of the matter in Chapter 4). The same argument can be pressed about the conditional in approaches that try to handle the Curry Paradox. Tarskianists have generally recognized this as a problem deserving an answer, and there are well-explored ways to answer it (for an example, see Priest, 2006, chs. 4–5).

However, the real problem with having to modify the operational rules is a different one entirely, whether one is changing the subject or not. The Tarskianists deal with the paradoxes piecemeal, missing the general features that allow the paradoxes to arise in the first place. But it seems that the paradoxes run deeper than any particular vocabulary. Tinkering with negation or conditional rules might prevent paradoxes involving negations and conditionals from arising, but it does not get to grips with the general phenomenon.

Two examples can reinforce this impression. One example is Curry-like paradoxes of validity, which will be discussed in the next section; the only thing we will say here is that no negation or conditional connective occurs in the validity Curry argument, so the Tarskianists' tweaks to negation and implication are beside the point here.

The other example is the Hinnion–Libert Paradox. The necessary resources for it are Reflexivity, Contraction, Cut and the following three groups of additional rules:

Membership

$$(\in L)\ \frac{\Gamma, A(a) \Rightarrow \Delta}{\Gamma, a \in \{x : A(x)\} \Rightarrow \Delta} \qquad \frac{A, \Gamma \Rightarrow A(a), \Delta}{\Gamma \Rightarrow a \in \{x : A(x)\}, \Delta}\ (\in R)$$

Extensionality

$$\frac{\Gamma, x \in a \Rightarrow x \in b, \Delta \qquad \Gamma, x \in b \Rightarrow x \in a, \Delta}{\Gamma \Rightarrow a = b, \Delta}$$

Identity

$$(=\mathrm{L})\ \frac{\Gamma, A(a) \Rightarrow \ \Delta}{\Gamma, a = b, A(b) \Rightarrow \ \Delta} \qquad \frac{\Gamma \Rightarrow A(a), \Delta}{\Gamma, a = b \Rightarrow A(b), \Delta}\ (=\mathrm{L})$$

$$\frac{\Gamma, Xa \Rightarrow Xb, \Delta \qquad \Gamma, Xb \Rightarrow Xa, \Delta}{\Gamma \Rightarrow a = b, \Delta}\ (=\mathrm{R})$$

The first of these groups of rules captures part of the expected behaviour of set membership in a naive set theory: the claim that a is a member of the set of ϕs can stand in for the claim that a is ϕ, either as a premise or as a conclusion. The rule in the second group ensures that the naive theory is a naive theory of sets, by allowing us to derive the claim that $a = b$ from a and b having the same members. The final group of rules captures part of the expected behaviour of '=': when drawn on as a premise, it allows substitution of one term for another in a premise.

From these resources, consideration of the Hinnion–Libert set $\{x : \{y : x \in x\} = \{y : \bot\}\}$ allows for a derivation of $\bot$. (A proof can be found in Restall (2013).) This paradox is in many ways just like a familiar (biconditional) Curry Paradox. But since it uses no connectives, in particular no conditionals and no negations, Tarskianist approaches to paradox must deal with it separately yet again, blaming at least one of the four additional rules. Of course they are welcome to do so; but again the impression is that they are missing the point: when each new paradox resembling the old ones must be blocked separately, it certainly gives the appearance that one is missing a general phenomenon.

Although we have presented the point so far as a problem for Tarskianist approaches that preserve transparent truth, it is equally telling against approaches to the truth-based paradoxes that work by giving up transparent truth, such as those explored by Maudlin (2004), Halbach (2011) and Scharp (2013). After all, neither validity Curry nor the Hinnion–Libert paradox (nor any of the more familiar set-theoretic paradoxes) has anything at all to do with truth; playing with our theory of truth may address the Liar and Curry Paradoxes, but it also misses the broader phenomenon.

Non-Tarskianist approaches take a different tack entirely, and one that seems to get at the root of the trouble. By blocking either Reflexivity, Contraction or Cut in their approach to the Liar, the non-Tarskianist have already blocked the problematic derivations from validity Curry or the Hinnion–Libert set, and no additional work is needed.

Rather than rushing from paradox to paradox making *ad hoc* modifications, these non-Tarskianist approaches seem to grapple with the paradoxes where they come from: the very basic features of argumentation. In this way, they can avoid having to worry about rules governing particular pieces of vocabulary; in a single fell swoop they address Liars, Curries, Validity Curries, Hinnion–Liberts, and so on, with the hope, of course, that no structural rule alone suffices to give rise to paradoxes.

Validity Curry

As mentioned in the previous section, a number of authors, including Whittle (2004), Field (2008), Shapiro (2011) and Beall and Murzi (2013), have put forward some arguments for the claim that logical validity is afflicted by paradoxes similar to the more well-known paradoxes that plague notions such as truth or set. But the idea that logical validity gives rise to paradoxes is in fact as old as Pseudo-Scotus. One version of the problem he discussed goes as follows. Consider the argument 'This argument is valid. Therefore, one plus one equals twenty six.' This argument is either valid or not. Suppose it is not. Then (there is a case such that) the premise has to be true and the conclusion untrue. But if the premise is true, the argument is valid, yet the conclusion is untrue, which goes against the definition of logical validity, so the premise cannot be true. But if it cannot be true, the argument is valid, for there would be no case in which the premise is true and the conclusion is false. Therefore, if it is invalid, it is valid. Suppose then that the argument is valid. If that is so, the premise is true. However, the conclusion is untrue. Then the argument is not valid. Therefore, if it is valid, it is invalid.

In the above version, the premise states the validity of the argument, while the conclusion is some suitable necessary falsehood. There is another version in which the premise is a suitable necessary truth and the conclusion states the invalidity of the argument. Consider the argument 'One plus one equals two. Therefore, this argument is invalid.' Again, the argument is either valid or invalid. If the argument is valid, then it cannot be that the premises are true and the conclusion is false. But, *ex hypothesi*, the argument is valid, so the conclusion is false. Then the premise should be false as well. But it is necessarily true, so the argument is not valid. Then, if it is valid, it is invalid. Suppose now that it is invalid. Then (there is a case such that) the premise is true and the conclusion is false. But by hypothesis the conclusion is not false, so the conclusion must be true as well, which it is. Then the argument is valid. Therefore, if it is invalid, it is valid. This paradox of logical validity serves to illustrate that the notion has been considered problematic long before contemporary logic and its formal techniques; we will not examine it further here, but the interested reader can consult Mates (1965), Read (1979) and Priest and Routley (1982).

More recently, worries about logical validity have been expressed through the investigation of the properties of a predicate '$Val(x, y)$', added to Peano arithmetic, PA, that holds for the Gödel code $\langle\phi\rangle$ of ϕ and the Gödel code $\langle\psi\rangle$ of ψ (in that order) if and only if the argument whose sole premise is ϕ and whose conclusion is ψ is logically valid. Of course, merely adding such a predicate to the language of arithmetic causes no more problems than merely adding a new predicate '$T(x)$' for truth does. Problems seem to arise once we supplement the axioms and rules of arithmetic with plausible rules for '$Val(x, y)$'. First, we have an *introduction rule* for '$Val(x, y)$':

$$I\,Val\text{: For any formulas } \phi \text{ and } \psi\text{, if } \phi \Rightarrow \psi \text{ then} \Rightarrow Val(\langle\phi\rangle, \langle\psi\rangle)$$

In short, *I Val* codifies the natural thought that if we have a proof of ψ from ϕ, then the argument with ϕ as premise and ψ as conclusion is valid.

With introductions out of the way, we also need something akin to an elimination rule for '$Val(x, y)$', and this is provided by:

$$E\,Val\text{: For any formulas } \phi \text{ and } \psi, \Rightarrow Val(\langle\phi\rangle, \langle\psi\rangle) \rightarrow (\phi \rightarrow \psi)$$

In short, *E Val* codifies the thought that validity preserves truth.

These are not the only possible, or only plausible, rules for the validity predicate. For our purposes here, however, these two rules suffice.[2]

We are now in a position to formulate the purported paradox of logical validity. First, we apply the Gödelian diagonalization lemma to the predicate '$Val(x, \bot)$' to obtain a sentence 'C' such that '$C \leftrightarrow Val(\langle C\rangle, \langle\bot\rangle)$' is a

[2] Some apparently plausible rules are in fact unacceptable, on pain of triviality. For example, an apparent version of the deduction property does not hold:

$$\text{For any formulas } \phi \text{ and } \psi\text{, and set of formulas } \Gamma\text{, if } \Gamma, \phi \Rightarrow \psi \text{ then } \Gamma \Rightarrow Val(\langle\phi\rangle, \langle\psi\rangle)$$

theorem. We can then, using arithmetic, *I Val* and *E Val*, derive a paradox along lines similar to the reasoning underlying the Curry Paradox:

1. C	[Definability of C]
2. $Val(\langle C\rangle, \langle\bot\rangle)$	[1, diagonalization]
3. $C \to \bot$	[2, *E Val*]
4. $\bot$	[1, 3, detachment]
5. $Val(\langle C\rangle, \langle\bot\rangle)$	[1–4, *I Val*]
6. $C \to \bot$	[5, *E Val*]
7. C	[5, diagonalization]
8. $\bot$	[6, 7, detachment]

On this paradox of validity involving a special predicate akin to the truth predicate there are roughly two sides. On the one hand, there are those who take it to be a genuine paradox and try to adapt some of the solutions to other paradoxes of self-reference to this case. On the other hand, there are those who consider this particular reasoning not to be paradoxical at all, because there is no reason to suppose that the rules for '$Val(x, y)$' are logically valid themselves; cf. Ketland (2012), Cook (2014).

Another way of expressing worries about logical validity is that of Field (2008) and Beall (2009), who have argued that logical consequence cannot be defined in terms of truth-preservation on pain of triviality. Suppose one wants a proof of the following biconditional:

> The conclusion of an argument is a logical consequence of the premises (in symbols, $P_1, \ldots, P_n \Rightarrow C$) if and only if the truth of the premises is preserved to the conclusion (in symbols, $(T\langle P_1\rangle \wedge \ldots \wedge T\langle P_n\rangle) \to T\langle C\rangle$).

Seemingly, the proof is straightforward. The validity of the inference from $P_1, \ldots, P_n$ to C is equivalent to the validity of the inference from

Let us suppose that ϕ stands for 'There is exactly one object' and ψ for 'There are exactly ten objects'. Then ϕ and ψ entail a contradiction:

$$\phi, \psi \Rightarrow \bot$$

It does not follow, however, that the claim that there is exactly one object entails that the argument whose premise is that there are exactly ten objects and whose conclusion is a contradiction is valid – that is:

$$\phi \not\Rightarrow Val(\langle\psi\rangle, \langle\bot\rangle)$$

$T\langle P_1\rangle, \ldots, T\langle P_n\rangle$ to $T\langle C\rangle$, by the usual truth rules. That in turn is equivalent to the validity of the inference from $T\langle P_1\rangle \wedge \ldots \wedge T\langle P_n\rangle$ to $T\langle C\rangle$, by the usual rules for conjunction. And that in turn is equivalent to the validity of the sentence $(T\langle P_1\rangle \wedge \ldots \wedge T\langle P_n\rangle) \rightarrow T\langle C\rangle$, by the usual rules for the conditional. But the validity of a sentence is but a strict form of truth, so this last step is just the claim that the inference preserves truth.

No matter how persuasive this argument looks, it turns on principles that cannot be jointly accepted. In particular, subscribing both to the truth rules employed in the first step of the argument and to the rules for the conditional employed in the last step, leads to triviality, as we learned from the Curry Paradox. In this case, no rule for a validity predicate can be blamed to fail to be logically valid, but every one of the different ways to resolve the Curry Paradox also undermines one or another step in the argument that validity is to be identified with truth-preservation. See Shapiro and Murzi (2015) for further discussion.

Acceptance, Rejection and Beyond

In recent years, there has been a growing tendency to define logical validity by appealing to concepts of cognitive states like *acceptance* and *rejection* (or of their linguistic expressions, *assertion* and *denial*), guided mainly by the idea that it would be irrational to accept the premises of a logically valid argument while rejecting the conclusion. But this joint acceptance of the premises and rejection of the conclusion can be avoided in several ways, each of them giving rise to different notions of logical consequence.

The obvious statement of logical validity in terms of acceptance and rejection can be obtained by simply substituting 'truth' uniformly with 'acceptance'. Thus, an argument is (generalized) Tarskian-valid if and only if, for every case the premises are *accepted*, the conclusions are also *accepted*. Equivalently, if the conclusions are *not accepted*, at least some of the premises are *not accepted* too. The relevant dialogical property is *acceptance*, so there is only one property forwards-preserved. That property determines another property, *non-acceptance*, which makes the collections of properties mutually exclusive and collectively exhaustive.

But a more "tolerant", so to speak, notion of logical consequence can be defined as follows:

Definition: An argument is *P-valid* if and only if, for every case the premises are *accepted*, the conclusions are *not rejected*. Equivalently, if the conclusions are *rejected* in some case, at least some of the premises are *not accepted*.

The relevant dialogical properties are those that count as *non-rejection*, since any of them can be preserved from *acceptance*. Again, *acceptance* and rejection are treated as mutually exclusive, but probably not collectively exhaustive, as there is the option of, say, hesitating; *non-acceptance* and *non-rejection* are collectively exhaustive, but probably not mutually exclusive: hesitation is not acceptance, but it is not rejection either.

However, with the possibility of tolerance usually comes the possibility of strictness, so a "stricter" notion of logical consequence can be defined as follows:

Definition: An argument is *Q-valid* if and only if, for every case the premises are *non-rejected*, the conclusions are *accepted*. Equivalently, if the conclusions are *non-accepted* in some case, at least some of the premises are *rejected*.

The relevant dialogical properties are those that count as *non-rejection*, but there is only one of them that has to be forwards-preserved: *acceptance*. As we have mentioned, *acceptance* and *rejection* might be mutually exclusive, but probably not collectively exhaustive; and *non-acceptance* and *non-rejection* might be collectively exhaustive, but probably not mutually exclusive.

These notions of logical consequence can be cashed out in the usual model-theoretic terms simply by giving up the requisite that the logical values used to define logical consequence need to be bi-partitioned.

Let us consider two of those strongly non-Tarskian notions of consequence, Malinowski's Q-consequence ('Q' for 'Quasi'; see Malinowski, 1990a,b) and Frankowski's P-consequence ('P' for 'Plausible'; see Frankowski, 2004a,b).

- Q-*consequence*. ϕ is a logical Q-consequence of premises Γ, in symbols $\Gamma \models^{Q} \phi$, if and only if any case in which each premise in Γ is not antidesignated is also a case in which ϕ is designated. Or, equivalently, there is no case in which each premise in Γ is not antidesignated, but ϕ fails to be designated.
- P-*consequence*. ϕ is a logical P-consequence of premises Γ, in symbols $\Gamma \models^{P} \phi$, if and only if any case in which each premise in Γ is designated is also a case in which ϕ is not antidesignated. Or, equivalently, there is

no case in which each premise in Γ is designated, but ϕ fails to be not antidesignated.

Suppose that there are three truth-values, $\top$, μ and $\bot$, with the order $\bot < \mu < \top$. Take $\top$ as the only designated value, $\bot$ as the only antidesignated value and μ as neither designated nor antidesignated. Suppose that $\mathfrak{v}(\phi) = \mu$. Then ϕ is not a logical Q-consequence of ϕ, because Q-consequence requires that if premises are not antidesignated, conclusions must be designated, which is not the case in this example. Q-consequence is not reflexive.

Take the same truth-values with the same order and the same partition. Suppose now that $\mathfrak{v}(\phi) = \top$, $\mathfrak{v}(\psi) = \mu$ and $\mathfrak{v}(\rho) = \bot$. Thus even if ψ is a logical P-consequence of ϕ and ρ is a logical P-consequence of ψ, ρ is not a P-consequence of ϕ, because P-consequence requires that if premises are designated, conclusions must be not antidesignated, which is not the case in this example. P-consequence is not transitive.

More Abstract Notions of Logical Consequence

The strongly non-Tarskian notions of consequence above demand *necessary preservation of non-antidesignatedness*. Clearly, designatedness is a case of non-antidesignatedness, so (generalized) Tarskian consequence is a case of a more general notion which also encompasses Q-consequence and P-consequence. For example, Q-consequence deals with preservation of non-antidesignated values but in such a strong way that it rather forces passing from non-antidesignated values to designated values. Similarly, P-consequence is preservation of non-antidesignated values but in such a weak way that it allows passing from designated values to some non-designated values (but never from designated to antidesignated values!). Let us call this notion of consequence *TMF-consequence* (for Tarski, Malinowski and Frankowski) and define it as follows:

Definition: ϕ is a logical *TMF-consequence* from premises Γ, in symbols $\Gamma \models^{TMF} \phi$, if and only if any case in which each premise in Γ is not antidesignated is also a case in which ϕ is not antidesignated. Or, equivalently, there is no case in which each premise in Γ is not antidesignated, but ϕ fails to be antidesignated. Again equivalently, if there is a case in which ϕ is antidesignated, then at least one premise in Γ is also antidesignated. In short, in the non-Tarskian notions of consequence the non-antidesignatedness of premises necessitates the non-antidesignatedness of conclusions.

Note that TMF-consequence has the following features:

(TMF1) It preserves a certain (kind of) value.

(TMF2) This preservation has a *direction* ("forwards"): it is from premises to conclusions.

(TMF3) The preservation in the other direction ("backwards") does not alter the logical consequences already determined by preservation-forward, so it can be dismissed.

Let's say that a relation R *backwards-preserves* a property ψ if and only if for any x, y such that Rxy, if y has the property ψ, x has it too. If the property ϕ forwards-preserved and the property ψ backwards-preserved are collectively exhaustive and mutually exclusive with respect to the properties that x and y might have, then forwards-preservation and backwards-preservation coincide. Consider the forwards-preservation of designated values (D^+-conseqence) and the backwards-preservation of antidesignated values (D^--consequence) as we defined them in Chapter 1. D^+- and D^--consequence distinguish directions in a way that forward-preservation and backward-preservation might not coincide. When the arrangement of values is such that $D^+ \cup D^- \neq \mathcal{V}$ (the designated and antidesignated values are not collectively exhaustive) or $D^+ \cap D^- \neq \emptyset$ (the designated and the antidesignated values are not mutually exclusive), D^+-consequence and D^--consequence do not coincide. Suppose, as above, that there are three truth-values, $\top$, μ and $\bot$, with the order $\bot < \mu < \top$ and with $\top$ as the only designated value, $\bot$ as the only antidesignated value and μ as neither designated nor antidesignated. Suppose pretty standard truth-conditions for conjunctions and conditionals, say $\mathfrak{v}(\phi \wedge \psi) = \inf(\mathfrak{v}(\phi), \mathfrak{v}(\psi))$ and $\mathfrak{v}(\phi \rightarrow \psi) = \top$ iff $\mathfrak{v}(\phi) = \inf(\mathfrak{v}(\phi), \mathfrak{v}(\psi))$, otherwise $\mathfrak{v}(\phi \rightarrow \psi) = \mathfrak{v}(\psi)$. Then ψ is a logical D^+-consequence, but not a logical D^--consequence, of $\phi \wedge (\phi \rightarrow \psi)$.

This suggests a more abstract notion of logical consequence, where the direction is abstracted, which can be called ***WS**-consequence*[3] stating the

[3] For Heinrich **W**ansing and Yaroslav **S**hramko, who have studied most of the abstract notions of logical consequence presented here, but whose main contribution to this topic has been the explicit distinction between the directions of D^+- and D^--consequence. See for example Wansing and Shramko (2008).

following: let premises Γ and conclusion ϕ be called *sides* of the relation of logical consequence and V and V^* certain (kinds of) values. ϕ is a logical WS-consequence from premises Γ, in symbols $\Gamma \models^{WS} \phi$, if and only if, if one side has values in V, the other has values in V^*.

Béziau's notion of logical consequence, which is axiomatically empty, results from WS-logical consequence when even the nature of the V_i is left unspecified. "Values" still have many "material" connotations, they have a very specific nature, so it would be better to say that they are structures of whatever kind. This does not mean that the last bit of logicality, in the sense of the core tradition, is lost. This rather means that the last step into abstraction could be taken if a better cartography of logicality is provided in a way that makes sense of both what we have called the core tradition in Chapter 1 and – what Béziau considers to be – the continuous journey of logic towards more abstraction.

To sum up: Traditionally, logics built upon the (generalized) Tarskian notion of logical consequence allow for variation in the cases in which premises and conclusions are evaluated, thus one obtains different particular logics. That is Beall and Restall's logical pluralism. It keeps the ideas that designated and antidesignated values form collectively exhaustive and mutually exclusive collections of values; that truth is the only designated value and that it is the value that must be preserved; that this preservation has a direction, namely forwards-preservation (i.e. from premises to conclusions) and that backwards-preservation yields the same results as forwards-preservation. An attempt to generalize this allows other values than truth to be designated and keeps the rest of elements as before. This is still Tarskian (at least generalized, i.e. reflexive and transitive) consequence nonetheless. Further generalization comes from the previous step yet allowing designated and antidesignated values to be either not collectively exhaustive or not mutually exclusive collections of values and allowing for different types of values to be preserved (only demanding that they are not antidesignated). These notions of consequence are no longer Tarskian, since they may be not reflexive (Q-consequence) or not transitive (P-consequence). WS-consequence abstracts the direction, kind of preservation and also the kind of value preserved: logical WS-consequence is a relation between values, where the order of the relata is undetermined. A bare logical structure is obtained when all the components of WS-consequence are left merely as mathematical entities.

Thus, logical consequence can be regarded as an octuple $LC = \langle \mathcal{V}; V_1, \ldots, V_n; \Vdash; V^P; V^C; CON; \delta; \mathcal{I} \rangle$ where $\mathcal{V}$ is a collection of structures that will serve as values; $V_1, \ldots, V_n$ are kinds of values such that each $V_i \subseteq \mathcal{V}$ and $V_1 \cup \ldots \cup V_n = \mathcal{V}$; V^P and V^C are the values for the sides of the relation (what are usually called 'premises' and 'conclusions', respectively); $\Vdash$ is a relation on $\mathcal{V}$; CON is the condition of connection in such relation, stating that if one side of the relation has a value of the kind V_i then the other has a value of the kind V_j; δ is the specification of the direction of this connection; finally, $\mathcal{I}$ is the family of indices at which values for the sides hold or otherwise. The notions of logical consequence discussed above appear naturally as specifications on the parameters of this octuple.

As we have seen in Chapter 7, Beall and Restall (2006) have argued that non-reflexive or non-transitive notions of logical consequence are not real logics, that they can be formalizing interesting phenomena in argumentation but not logical consequence, so there is a legitimate concern that the generalizations considered here might not be logics, especially when logical consequence is expected to satisfy certain features like normativity or formality, introduced in Chapter 1. Although the issue is highly contentious, we will indicate that these kinds of logics are not so easily excludable from the realm of logic proper.

Consider Q-consequence and P-consequence. A good signal that we are not very far from what is commonly called 'logic' is that, under minimal classical constraints on the structure of truth-values, these notions of consequence are indistinguishable from the Tarskian one. If there are only two truth-values, *true* and *false* with their usual order, the collections of designated and antidesignated values can exhaust all the possible values, hence *designated* = *not antidesignated* and *not designated* = *antidesignated*. But if this were merely a feature of classical logic, surely Q-consequence and P-consequence would have emerged earlier than they actually did. However, these notions of consequence collapse if the collections of designated and antidesignated values are supposed to be mutually exclusive and collectively exhaustive with respect to the total collection of values given, as is assumed in almost every known logic.

At least technically, strongly non-Tarskian notions of consequence such as Q-consequence and P-consequence are as legitimate as non-classical logics are. More elaborate answers could be given along the lines of non-monotonic logics. For example, Q-consequence would serve to "jump" to conclusions more certain than the premises, and P-consequence would allow us to jump to conclusions less certain than the premises.

There is no *prima facie* reason as to why it is possible to do without monotonicity but not without reflexivity or transitivity: as properties of a relation of logical consequence they seem to be *pari passu*. Beall and Restall's claim that they are not on equal footing "because preservation of designated values (from premises to conclusions) is a reflexive and transitive relation" begs the question: non-reflexive and non-transitive logics are precisely asking for ways of logically connecting premises and conclusions other than preservation of truth or, more generally, of designated values.

Now, granting that there is a core tradition in logic which demands that logic is at least necessary, normative and formal, the burden is on the proponent of strongly non-Tarskian notions of logical consequence to show that they somehow still belong to that tradition. More exactly, they have to show either that the core tradition underlying truth-preservation is wrong or that it makes room for non-reflexivity and non-transitivity. We think a strong case can be made for the latter option. Let us discuss *necessity* first. As we have seen, the strongly non-Tarskian notions of logical consequence above demand *necessary preservation of non-antidesignatedness*, that is, the non-antidesignatedness of the premises necessitates the non-antidesignatedness of the conclusions. Even in the case of WS-logical consequence, the values – whatever they are – on one side of the relation – whatever it is – necessitate a certain value on the other.

Secondly, consider *normativity* and the applicability to the evaluation of arguments in daily life and sciences. It is easy to see that, in spite of appearances, we are still in the business of logic. As we have seen, if one uses cognitive states like acceptance and rejection (or their linguistic expressions, assertion and denial) to define validity, these different notions of consequence arise almost naturally:

- An argument is TMF-valid if and only if, if the premises are not rejected, then conclusions are also not rejected. Equivalently, if the conclusions are rejected, premises are rejected too. The relevant dialogical properties are those that count as *not rejected*, so there is more than one property that could be forwards-preserved. The preservation of these properties is required, but the direction is not important since *not rejected* determines another property, whose name depends on what the components of the first property are taken to be that make the collections of properties mutually exclusive (but probably not collectively exhaustive). Now, in this case, not rejecting the premises but rejecting the conclusions might be regarded as irrational in those contexts where TMF-logical consequence applies.

- An argument is (generalized) Tarskian-valid if and only if, if the premises are accepted, then the conclusions are also accepted. Equivalently, if the conclusions are not-accepted, the premises are not-accepted either. The relevant dialogical property is *accepted*, so there is only one property forwards-preserved. That property determines a property, *not accepted*, which makes the collections of properties mutually exclusive and collectively exhaustive.

This is still so when one moves towards WS-logical consequence; the normative aspect can be inherited form the other notions:

- An argument is WS-valid if and only if, if one side of the argument has a dialogical property, then the other also has it. No dialogical property is especially relevant nor are properties required to have further structure, such as being mutually exclusive; neither preservation of dialogical properties nor a specific direction in which they are connected is required, it's only required that there is some connection between them. Not respecting the connection might be regarded as irrational in the relevant circumstances.
- An argument is accepted-valid if and only if, if the premises are accepted, then the conclusions are also accepted. The relevant dialogical property to be preserved is *accepted*; this preservation has a preferred direction (forwards).
- An argument is rejected-valid if and only if, if the conclusions are rejected, then the premises are also rejected. The relevant dialogical property that has to be preserved is *rejected*; this preservation has a preferred direction (backwards).

Accepted-validity and rejected-validity need not coincide, for acceptance and rejection, even if mutually exclusive, are not collectively exhaustive. Similar variations can be obtained using not only dialogical, but also probabilistic, psychological, epistemic or whatnot specifications of consequence, for example:

- An argument is inconceivable-valid if and only if, in any situation that is inconceivable for the conclusions to be true, it is also inconceivable for the premises to be true. The relevant ("psychological") property that has to be preserved is *inconceivability*; this preservation has a preferred direction (backwards).
- An argument is believed-valid if and only if, for every case in which the premises are believed, the conclusions are believed too. The relevant ("epistemic") property to be preserved is *believed*; this preservation has a preferred direction (forwards).

Other variations are left to the reader; and in general normativity operates since it could be irrational not to respect the required connection between premises and conclusions in the contexts where each notion of logical consequence operates.

About *formality*: There is an enormous lack of precision on this matter, but it is fairly innocuous given that the non-Tarskian notions of logical consequence are generalizations of Tarskian logical consequence, as we have shown. The meanings of no other terms than those accepted by the (generalized) notion of Tarskian consequence are in play here, so there is no threat to formality if the original notion is already formal. This applies even to the notions of WS-logical consequence and of Béziau's logical structure.[4]

As to other elements of the core tradition, formality is kept since it is an abstraction from the previous notions, so WS-consequence is formal if the previous notions are, and necessity is also kept since a (kind of) value on one of the relata necessitates some (kind of) value in the other.

Questions

1. Reading $\Gamma \Rightarrow \Delta$ as 'in every case where Γ is *accepted*, Γ must not (all) be *rejected*', revisit the sequent calculus in Chapter 2 and see which rules begin to look suspicious.

[4] Greg Restall has stressed the importance of these cognitive, normative notions for the definition of logical consequence and theories, but his approach is different. Rather than studying separate notions of consequence stressing either of these cognitive notions, Restall studies interactions of them in a single notion of logical consequence to give a more comprehensive notion of *(bi)theory* (Restall, 2013).

2. On p. 207 we claimed that consideration of the set $\{x : \{y : x \in x\} = \{y : \bot\}\}$ allows for a derivation of $\bot$. Show how. (You might want to consult Restall (2013) for inspiration.)
3. In a note on p. 205 we claimed that the Curry derivation didn't rely on adding the special two-way conditional rule to our system. Convince yourself this is in fact true by either re-deriving $\phi \Rightarrow \psi$ in the sequent calculus from Chapter 2 in the presence of a Curry sentence or by showing that every instance of $\to I/E$ is derivable in that system.

10 The Place of Logic in Science

'Logic' is the name of a philosophical subdiscipline, and – as you learned already at the very beginning of the book – it is also the name of a subdiscipline of mathematics, a theory of implication, the abstract network of actual implications, and a capacity or disposition to make certain inferences. In this chapter we want to discuss "the place of logic in science". This title inherits the ambiguity of 'logic'. This time we will not disambiguate though. We will indeed look at the place of logic in science under different meanings of 'logic'. For the most part, however, our focus will be on logical theory, *logica docens*, and the role it plays in other disciplines and how it relates to the theories in those other disciplines.

Logic and the Empirical Sciences

If we think of logic as a theory of logical consequence then there is one quite obvious relation between that theory and the theories of other sciences. Every science must be interested in whether its theoretical account is consistent, and whether it accords with the observable facts. The former is clearly a matter of logical analysis. But the latter is also partly a matter of logical analysis, in so far as it requires the logical derivation of empirical statements which can be put to test in, for example, an experiment.

As you probably know from introductions to philosophy of science (or the brief discussion in Chapter 6), empirical theories by themselves typically do not logically imply specific predictions. Such predictions are *derived* from a theory together with certain background assumptions and auxiliary hypotheses. To take a toy example, the theory 'All ravens are black' does not imply that the bird you are looking at appears black to you. It only does so under the background assumption that you are looking at a (prototypical) raven, the auxiliary hypothesis that normal human vision under circumstances *X* is a reliable way to detect the actual colour of objects looked at, that you possess normal human vision, that the observation is taking place under circumstances *X*, etc. However, whatever other theories or assumptions might be needed to derive empirical consequences from a theory, without using logic in some sense, you won't be able to derive anything.

Whether this establishes a foundational role for logic *as a theory* in the sciences is a topic that we will postpone until the end of this chapter. What we will discuss in the next few sections are cases in which logical theory plays a somewhat different role, cases in which logical theory is directly relevant for the projects of other empirical sciences.

We begin with an example at the intersection of logic/philosophy of language and linguistics, the study of *formal semantics*.

Formal Semantics

At the basis of formal semantics is Richard Montague's insight that the relation between syntax and semantics in a natural language such as English could be viewed as not essentially different from the relation between syntax and semantics in a formal language such as the language of first-order predicate logic.

The semantics of formal language is, as you know from Chapter 1, usually provided by model theory, a mathematical theory, in particular a branch of mathematical logic. Why believe that logic is of any help for the study of natural language semantics? There are at least two good reasons for that. The first is that – at least since Gottlob Frege – logic is conceived as the science that studies the laws of truth. But then, truth and meaning have an obvious connection in what Max J. Cresswell called the *Most Certain Principle* (Cresswell, 1973):

Cresswell's Most Certain Principle For two sentences α and β, if in some possible situation α is true and β is false, α and β must have different meanings.

A second reason is that the task (or, in any case, one task) of a semantic theory is to correctly predict meaning relations between expressions. But, arguably, all meaning relations are (or are reducible to) logical relations. Thus, e.g. hyponymy is (nothing but) schematic implication: *green* HYP *coloured* because ⌜x is green⌝ implies ⌜x is coloured⌝, whatever x is; incompatibility is also reducible to implication (and negation): *green* INCOMP *red* because ⌜x is green⌝ implies ⌜x is not red⌝, i.e. the negation of ⌜x is red⌝. Thus understood, a logical relation among linguistic expressions is one that can be reduced to the notion of logical (or analytic) implication. Formal semantics, in this sense, can be understood as the attempt to describe all relevant semantic relations in terms of logical relations (Zimmermann and Sternefeld, 2013).

At the heart of almost all work in formal semantics lies another Montagovian idea (usually traced back to Gottlob Frege), namely the idea that the semantics of a natural language must be compositional. The so-called "principle of compositionality" is thus often regarded as a fundamental principle of formal semantics:

The Principle of Compositionality The meaning of an expression is uniquely determined by the meanings of its parts and their syntactical combination.

It is true that the Principle of Compositionality helps to explain how formal semantics usually works. A formal semantics considers fragments of natural language which is such that from a base class of "basic expressions" a class of grammatical "complex expressions" can be generated by the application of syntactical rules (in Figure 10.1: SR_1, SR_2, …) of the language. The task of a formal semantics is then to provide an interpretation for the expressions of the base class and a set of composition rules that mirror the syntactical rules and determine the semantical values of the complex expressions from the semantical values of their parts (in Figure 10.1: CR_1, CR_2, …).

Aside from providing a neat way to explain the approach of formal semantical analysis, the Principle of Compositionality has a rather unclear

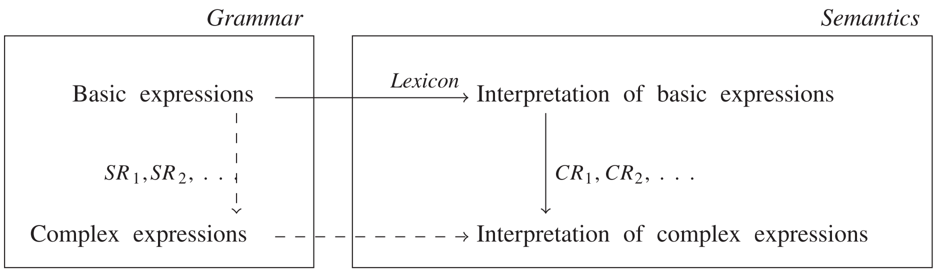

Figure 10.1 Compositional semantics (adapted from Löbner, 2013)

status within formal semantics. It is sometimes (and also since Montague) invoked as an explanatory constraint on any natural language semantics to make explainable the fact that finite human beings can understand languages with infinitely many expressions with different meanings. In this sense, the principle seems to be of explanatory value, and hence empirical in nature. However, some formal semanticists claim that in the framework of a formal analysis of meaning the principle is actually trivial (and hence not empirical) and should best be regarded as a methodological principle. Of course, for the adoption of a methodological principle one would still like to hear reasons, but so far none have been provided (Cohnitz, 2005).

The principle is, in any case, not unproblematic, since it directly leads the semantical analysis into (what is called) "Mates' Trap" or "Mates' Puzzle" (Mates, 1952) if combined with Cresswell's Most Certain Principle and the rather uncontroversial assumption that synonymy is logical indiscernibility and thus weaker than identity. The trap is that the three assumptions taken together seem to contradict the fact that, for any two distinct expressions α and β, the first of the following sentences should be true whereas the second appears to be false:

1. It is possible for someone to believe that α is the same as α without believing that α is the same as β.
2. It is possible for someone to believe that α is the same as α without believing that α is the same as α.

By Cresswell's Most Certain Principle, sentences 1 and 2 must differ in meaning, and thus by the Principle of Compositionality, α and β also must have different meanings.

Some semanticists claim that this refutes the Principle of Compositionality (Pelletier, 1994); whereas others argue that it shows the limits of formal semantics. (It is, however, also arguable that the meaning difference between sentences 1 and 2 should be analysed in pragmatic rather than semantic terms (Stalnaker, 1999).)

Formal semantics usually proceeds in its compositional analysis in two steps. In the first step, expressions of the natural language are translated into a formal language (often some variant of predicate logic). Then the formal language is given a semantic interpretation within model theory.

In this procedure the formal language as such does not play any special role. It would be possible (and it is sometimes done) to make the

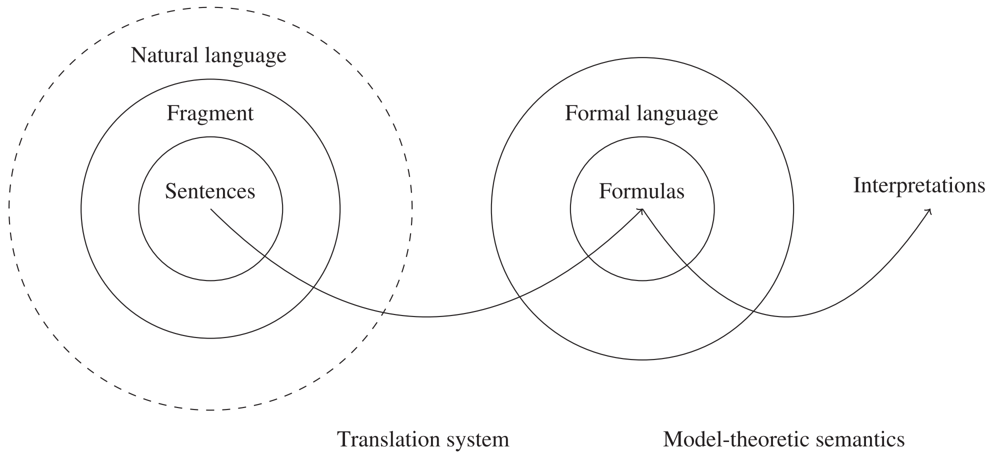

Figure 10.2 Formal semantics (adapted from Löbner, 2013)

model-theoretic meaning assignment directly to the expressions of the natural language without the "detour" through the formal language (and combine both steps of analysis into one). However, this would make the semantic evaluation less transparent and more difficult to parse. The meaning assignment to the formal language is, on the other hand, a formally determined procedure that could even be skipped; the interesting part of the analysis in practice thus really is the *translation* in such a two-step analysis (see Figure 10.2).

As indicated in Figure 10.2, formal semantics so far only deals with fragments of natural languages, and it is far from being in a position to deal with any complete natural language. However, formal semantics is a thriving field of research, making heavy use of formal tools developed within mathematical logic and model theory. As we have seen at several places in this book, research at the intersection of logic and natural language semantics is not only of relevance for linguistics, but often informs logical theorizing (as one example, remember the discussion of plural quantification and the Geach–Kaplan sentence in Chapter 3).

Cognitive Science

The next science we look at is a field of intensive interdisciplinary work. Cognitive science is the interdisciplinary study of the mind and of cognition in particular. It comprises approaches from philosophy, artificial intelligence, psychology, neuroscience, linguistics and anthropology. Obviously, one place in which cognitive science and logic meet is in the study of *human reasoning*.

We discussed the relation between logic and the study of human reasoning earlier, in two places in this book. In Chapter 5 we discussed *psychologism*, the idea that logic is really the study of human reasoning, and in Chapter 8 we discussed in what sense logic can play a normative role for human reasoning in its relation to a theory of theoretical rationality. In this chapter we will have a slightly different focus. We will discuss the role that logic (as a theory of logical consequence) can have for developing a *descriptive* or *explanatory* account of human reasoning.

Of course, in the heyday of psychologism, when logic was seen as a descriptive psychological theory, there was a clear sense in which logic was of relevance for psychology (because it *was* psychology). But after logic "emancipated" itself from psychology (in the development of modern mathematical logic), logic was also playing an explanatory function within psychology. As Stenning and van Lambalgen (2008) explain, cognitive science took off when adopting the information processing metaphor, consisting of the following three methodological assumptions:

1. Cognitive explanations must refer to models, conceived of as representational mechanisms
2. which function "in the same way" as the phenomena being represented
3. and which are capable of generating behavior and thoughts of various kinds. (Stenning and van Lambalgen, 2008, 8)

Logic played a double role in this scheme, as a "formal, symbolic, representation language", and as "an inference mechanism generating behavior and thoughts". At least normal adults reason logically, automatically as well as reflectively. This assumption, that humans use a domain-general logic in reasoning, has come under attack in psychology. According to Stenning and van Lambalgen (2008) this can be explained as a reaction to certain empirical results, such as, most prominently, the *Wason Selection Task* (aka the "four-card problem").

In 1966 the cognitive psychologist Peter Cathcart Wason devised a test that purportedly showed the irrationality or illogicality of human reasoning (Wason, 1968). He presented his test subjects with the task depicted in Figure 10.3.

Below is depicted a set of four cards, of which you can see only the exposed face but not the hidden back. On each card, there is a number on one of its sides and a letter on the other.
Also below there is a rule which applies only to the four cards. Your task is to decide which if any of the cards you *must* turn in order to decide if the rule is true. Don't turn unnecessary cards. Tick the cards you want to turn.

Rule: *If there is a vowel on one side, then there is an even number on the other side.*

Cards:

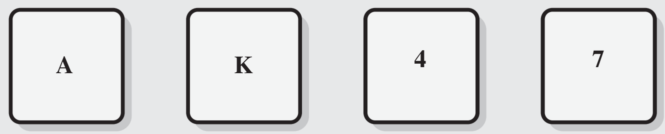

Figure 10.3 Wason Selection Task (from Stenning and van Lambalgen, 2008)

If we code 'there is a vowel on one side' as 'p' and 'there is an even number on the other side' as 'q', then test subjects typically pick cards in this test in accordance with the frequency displayed in Table 10.1.

Table 10.1 *Typical results for the Wason Selection Task (from Stenning and van Lambalgen, 2008, 46)*

p	p, q	$p, \neg q$	$p, q, \neg q$	misc.
35%	45%	5%	7%	8%

Wason assumed that the rule that was given in the task (*If there is a vowel on one side, then there is an even number on the other side*) is to be understood as a material conditional, '$p \rightarrow q$' of classical logic. Thus, the "correct" answer in this task should be to select (only) 'p' and '$\neg q$', which 95% of the test subjects failed to do.

Subsequent experiments showed that test subjects performed much better when the test did not involve such an abstract rule (in terms of vowels and numbers, for example), but was rather cast in terms of rules that test subjects were purportedly more familiar with. Thus, Johnson-Laird et al. (1972) showed that if you present test subjects with the rule 'If a letter is sealed, then it has a 50 lire stamp on it' and a series of four pictures showing the backs and fronts of letters (sealed or not, with or without a 50 lire stamp on it), almost everyone gets it "right".

Such results led to the speculation of evolutionary psychologists that humans don't reason with a domain-general logic, but instead have developed a domain-specific competence to check rules that have to do with social contracts, in order to able to efficiently detect "cheaters" (Cosmides, 1989; Cosmides and Tooby, 1992). But then logic, as a domain-general theory, seems useless for an explanation of actual human reasoning.

> The study of human reasoning has been dominated by the search for content-independent cognitive processes. Early research started from the premise that humans reason logically, that is, using the rules of inference of the propositional calculus. These rules of inference are content-independent: they generate only true conclusions from true premises, regardless of what the propositional content of the premise is. However, more than a decade of research has shown that people rarely reason according to these canons of formal logic. Moreover – and contrary to initial expectations – psychologists found that human reasoning is content-dependent: the subject matter one is asked to reason about seems to regulate how people reason. Nowhere is this seen more clearly than in experiments using the Wason selection task. (Cosmides, 1989, 191)

Instead of trying to describe human reasoning with the help of a domain-general logic, we should rather focus on domain-specific reasoning modules or "fast and frugal algorithms" that our evolutionary history has equipped us with (cf. also Gigerenzer, 2000).

However, this is not the only possible interpretation of the empirical results. In a very careful study, Stenning and van Lambalgen (2008) argue that we should distinguish in such tasks the *reasoning to an interpretation* and the *reasoning from an interpretation* of the linguistic item that supposedly expresses the rule in question.

Most of the psychological literature on the Wason Selection Task seems to assume that the rule to be evaluated by the test subject has an unambiguous logical form and that, moreover, this logical form is the same in the abstract and the "realistic" examples. But this can be disputed. As Stenning and van Lambalgen (2008) show, the "abstract" cases are best interpreted as "descriptive" conditionals, which then lead to several difficulties of interpretation for the test subject. The "realistic" cases, on the other hand, seem to be deontic conditionals, which are "easier" since they don't require the same interpretational effort from the test subjects.

In other words, there is a reasoning process that leads to an assignment of logical form, and a reasoning process that begins after the logical form is fixed. The empirical results do not show that the two processes can't be described by (some) domain-general logic. In fact, Stenning and van Lambalgen (2008) argue that the first process, that of reasoning *to* an interpretation, is best analysed by some non-monotonic logic, while the second process, that of reasoning *from* an interpretation, *may* sometimes well be described by classical logic (see also Stenning and van Lambalgen, 2011).

We should tie these considerations together with some of the topics that we discussed in Chapter 8. Following Stanovich (1999), Stenning and van Lambalgen (2008) distinguish between the following three types of rules:

- *Normative rules*: reasoning as it should be, ideally
- *Descriptive rules*: reasoning as it is actually practised
- *Prescriptive rules*: norms that result from taking into account our bounded rationality, i.e. computational limitations (due to the computational complexity of classical logic, and the even higher complexity of probability theory) and storage limitations (the impossibility of simultaneously representing all factors relevant to a computation, say, of a plan to achieve a given goal). (Stenning and van Lambalgen, 2008, 6)

Normative rules may comprise *modus tollens* or *Bayes' Theorem*, *descriptive rules* may comprise certain common fallacies (such as the *Base-Rate Fallacy*[1]),

[1] *Bayes' Theorem* is the name of the following equation:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Here A and B are events, and '$P(A)$' is the probability of (observing) event A (independently of B), and '$P(A \mid B)$' is the conditional probability of (observing) event A given that B occurs; it is assumed that $P(B) \neq 0$.

Bayes' Theorem can then be used to calculate such conditional probability. Assume that you tested a random person with a highly reliable drug test (the test is so reliable that it produces merely 1% false positives and 1% false negatives), and the test came out positive. How likely is it that the person tested indeed consumed the drug? You might think that this probability must be high, since your test is so reliable. However, as Bayes' Theorem reminds us, this ignores the *base rate*, the probability of someone using the drug in the first place. Let's assume that the drug doesn't have many users, only 0.5% of the people use it. Plugging these numbers into Bayes' Theorem shows that the conditional probability for your test subject to be a drug user, given that the test came out positive, is only about 33.2%. Ignoring the base rate and treating $P(A \mid B) = P(B \mid A)$ is called the *Base-Rate Fallacy*.

and *prescriptive rules* are best thought of as classically invalid principles that hold defeasibly under specific (common) circumstances.

Once these rules are conceptually distinguished, we can identify the different strategies for how to interpret the empirical data concerning human reasoning from a theory of rationality point of view (as discussed in Chapter 8).[2]

Panglossians consider human reasoning competence *and* performance to be actually correct. The apparent incorrectness (as it seems to show up in the Wason Selection Task, for example) can be explained away by clever choices in the task construal or the interpretation of logical connectives, etc.

Apologists on the other hand admit that humans follow merely subnormal prescriptive rules (because following the proper norms is computationally too demanding for us).

Meliorists hold that we fall short of even following the subnormal prescriptive rules.

Eliminativists finally hold that reasoning, in the sense of following any domain-general rules, doesn't really ever happen. We only follow domain-specific algorithms that have evolved under constraints of time and energy.

We have seen that, typically, evolutionary psychologists are Eliminativists. The position that Stenning and van Lambalgen (2008) defend is not Panglossian because they hold that, for example, once an interpretation of the task is fixed, subjects can make mistakes in reasoning from that interpretation. However, it is not completely Apologist or Meliorist either because they don't follow the normative/prescriptive distinction for rules. Instead, they are a certain type of Meliorist who hold that humans try to follow domain-relative norms of logic.

Logic and Mathematics

The relationship status between mathematics and logic is best described with Facebook's famous 'it's complicated'. Of course, contemporary logic is in part studied in mathematics departments and, even when it's studied by philosophers, it is studied with the help of mathematical tools. But what is the relationship between mathematics and logic?

[2] The following terminology is again adapted from Stanovich (1999).

Let us begin with the perhaps more prominent view that mathematics just reduces to logic. This view in philosophy of mathematics is called 'logicism'. The first serious attempt to achieve such a reduction was made by Gottlob Frege, who sketched his project in *Grundlagen der Arithmetik* (Frege, 1884) and then tried to carry it out in detail in *Grundgesetze der Arithmetik* (Frege, 1893). In order to understand in what sense Frege tried to reduce mathematics to logic, we need to clarify what he understood by 'mathematics', 'logic' and 'to reduce'.

The intended reduction aimed at showing that mathematics was analytic in an epistemological sense, i.e. that it consists only of general logical laws or definitions or of theorems that have proofs relying on such general logical laws or definitions only. 'Mathematics' for Frege meant arithmetic (as far as geometry was concerned, Frege followed Kant in believing that its truths are *synthetic*), and the logic that Frege used was a higher-order logic.

Frege managed to define *equinumerosity* in merely logical terms. The definition of 'equinumerous' says that two concepts are equinumerous if there is a 1–1 mapping of the objects falling under them. This allowed him to formulate "Hume's Principle" in exclusively logical terms:

Hume's Principle For any concepts F, G, the number of F is identical to the number of G if and only if F and G are equinumerous.

or in the notation of second-order logic (where '$\#xFx$' is a singular term, symbolizing 'the (cardinal) number of' the Fs'):

$$\forall F \forall G(\#xFx = \#xGx \leftrightarrow \exists R(\forall x(Fx \rightarrow \exists y \forall z(z = y \leftrightarrow (Gz \wedge Rxz))) \\ \wedge \forall x(Gx \rightarrow \exists y \forall z(z = y \leftrightarrow (Fz \wedge Rzx)))))$$

Indeed, Hume's Principle is sufficient for the derivation of the Peano–Dedekind postulates for the arithmetic of natural numbers (Tennant, 2015). Frege, however, tried to achieve more than that and base his theory of numbers in a general theory of extensions, which was shown by Bertrand Russell to lead to paradox.

Contemporary *neologicists* of the "Scottish School" follow Frege's project of epistemic foundationalism, explaining our knowledge of arithmetic (and perhaps all of classical mathematics) on Frege's *Context Principle*, abstraction principles and second-order logic (Ebert and Rossberg, 2007):

> Roughly speaking, the function of the context principle is to guarantee that mathematical singular terms indeed refer, and so refer to abstract objects. The theory of abstraction principles aims firstly, to introduce mathematical singular terms and secondly, to offer a "epistemically tractable" way of how a subject can come to know basic mathematical principles. Lastly, second-order logic is adopted in order to generate the theorems of mathematics. (Ebert and Rossberg, 2007, 35)

Whether such a programme indeed establishes that mathematics is reducible to logic (in some epistemological sense) is of course a matter of whether the ingredients of that epistemic reduction can be considered purely logical.

We already discussed in Chapter 3 the question of whether second-order logic should be considered logic proper. If second-order logic really is mathematics in "sheep's clothing" then neologicism might still be a very interesting project (showing that a certain part of mathematics is reducible to another part of mathematics), but not foundationalist in the intended sense.

A more pressing problem might seem to be whether *abstraction principles*, such as Hume's Principle above, violate the requirement to use only logical ingredients in the foundation of mathematics. After all, logic is supposed to have no ontology, but then how can purely logical principles lead to a theory of *numbers*?

Abstraction principles are meant to function as implicit definitions. They are stipulated as true, and introduce new expressions; in this case the expression 'is the (cardinal) number of'. However, how do we know that the right-hand side (the side to the right of the first occurrence of '$\leftrightarrow$') of such an abstraction principle is true? Ebert and Rossberg (2007) explain the reasoning of the neologicists as follows: Arguably, it is a logical truth that the instances of the concept *being non-self-identical* can be put into a one-to-one correspondence with themselves. This is trivially so, because there are no things that are not self-identical, which is held to be a logical truth. This gives us

$$(\#x(x \neq x) = \#x(x \neq x)) \leftrightarrow ((x \neq x) \simeq (x \neq x))$$

In this expression, we abbreviate 'there is a one-to-one correspondence between', with the symbol '$\simeq$'.

If we assume that Hume's Principle is true, and since the right-hand side of the previous equivalence is a logical truth, we can discharge that right part and get to

$$(\#x(x \neq x) = \#x(x \neq x))$$

Since, '$\#xFx$' is a singular term, we can existentially quantify into the expression above, and get the following existential statement:

$$\exists y(y = \#x(x \neq x))$$

If the quantifier is understood as existentially committing, knowledge of logic (in some sense), together with certain stipulative definitions, leads to the required ontology to build up a theory of numbers. Obviously, this story contains several '*if*s', but it is a very promising theory for explaining our knowledge of mathematics and of numbers in particular.

Is Metalogic Mathematics?

But as we said at the beginning of this section, the relationship between mathematics and logic is complicated. As we have discussed in Chapter 3, the metatheory of logic (model-theoretic semantics in particular) makes use of mathematical notions (such as sets). Doesn't that show that the foundations of logic (even those of first-order predicate logic) are actually mathematical? Note that this question is not only a matter of concern for a neologicist. For example, many nominalist programmes in the philosophy of mathematics rely on logical notions (such as consistency or logical entailment) when accounting for the purported ontological innocence of mathematics. However, if we can provide an account of these logical notions only by invoking mathematical objects, how can the nominalist account of mathematics then be said to have shown that mathematics is indeed dispensable (see Wilholt, 2006)?

One possibility is to take certain logical notions to be simply primitive. So, *logical consequence* is perhaps simply a primitive notion that is not in need of an account. This is the strategy of, for example, Field (1991). Some nominalists might find this move unacceptable; arguably the notions of logical consequence – of the technical kind that is required for nominalist programmes to succeed – is too far away from our intuitive understanding to be considered *primitive*:

> When beginning students are first told about logical possibility, logical consequence, etc., most of them seem to have some idea of what is meant, but consider how much their initial "intuitions" differ from our "refined" ones. The anti-realist owes us some account of how we plausibly could come to understand the notions in question (as applied here) as we in fact do, independent of our mathematics. Without this it is empty to use a word like "primitive" [. . .]. (Shapiro, 1993, 475)

An alternative strategy would be to show that metalogical notions, such as that of *logical consequence* can be explicated in nominalistically acceptable terms, quantifying over nominalistically acceptable entities (Rossberg and Cohnitz, 2009).

There is a further way in which mathematics could be considered foundational to logic, rather than the other way around. The logicist and neologicist programmes that we discussed above take logic to be fundamental and then try to account on that basis for mathematics. However, there is an influential tradition in philosophy of mathematics to see that relationship between logic and mathematics to be exactly reversed. Perhaps mathematics and mathematical reasoning should be considered to be fundamental and an account of logic should follow from that basis. This is the tradition known as "intuitionism" in the philosophy of mathematics.

Intuitionism in the Philosophy of Mathematics

The version of intuitionism in mathematics which supports this view on the relationship between mathematics and logic goes back to the work of the Dutch mathematician Luitzen Egbertus Jan Brouwer (1881–1966). Brouwer held a view similar to that of Immanuel Kant on which mathematics is founded on a type of "pure intuition". In Brouwer's philosophy, mathematics is founded on the pure intuition of inner time. This is to be understood not as a psychological theory, but rather as a psychologistic account of mathematics that is rooted in an *idealization* of mathematical thinking (Iemhoff, 2016). In any case, on this view mathematics is a "languageless activity" and thus separate from "mathematical language and hence from the phenomena of language described by theoretical logic" (Brouwer, 1981, 4).

Therefore logic is neither the foundation of, nor a constraint on, mathematics and mathematical reasoning. Mathematicians use language to communicate and exchange their mathematical ideas, but the ideas are themselves independent of the language used for their transmission. But then logic is a description of the patterns found in the linguistic communication of mathematical ideas. Logic, on this view, is not normative for thought. It's neither a normative theory nor normative in the Fregean sense of being a proper guide for reasoning due to the fact that it codifies the (descriptive) laws of truth. The source of mathematical innovation is the free mind of the creative subject. Logic might not yet codify what intuition may still come up with.

Intuitionistic mathematics is in several ways different from "classical" mathematics. It is a form of constructivism and a truly revisionary account of mathematics. Neither is all of classical analysis intuitionistically acceptable, nor is all of intuitionistic analysis classically acceptable. The logic that is lifted from intuitionistic mathematics is *intuitionistic logic*. We don't have the space to discuss it in any detail, but still want to provide the main idea behind it.

As we said, intuitionism is a kind of constructivism. Mathematical entities are constructed, not discovered, and so claims about them must be the result of (constructive) proofs. A logic that is to describe this reasoning will not contain the law of excluded middle, $(\phi \vee \neg\phi)$. Brouwer provided "weak counterexamples"[3] against this law. Consider the *Twin Primes Conjecture*, i.e. the conjecture that there are an infinite number of *twin primes*, primes of the form $(p, p+2)$. This conjecture that has not (yet) been proven in mathematics (example from Moschovakis (2015)). Thus, we have no proof, either for it or for its negation. Let x, y range over the natural numbers 0, 1, 2, ... and '$B(x)$' abbreviate the property expressed by the following claim in which the variable x is free: there is a y greater than x such that both y and $y+2$ are prime numbers, or, formally:

$$\exists y(y > x \wedge \text{Prime}(y) \wedge \text{Prime}(y+2))$$

We have no general method for deciding whether '$B(x)$' is true or false for arbitrary x since we have no proof of the conjecture, so '$\forall x(B(x) \vee \neg B(x))$' cannot be asserted in the present state of our knowledge. But if 'A' abbreviates

[3] These are considered "weak" counterexamples, because they merely show that the law of excluded middle has instances for which we don't have positive grounds, but don't strictly refute the law (see van Atten, 2017).

the statement '$\forall x B(x)$', then '$(A \vee \neg A)$' cannot be asserted because neither 'A' nor '$\neg A$' has yet been proved.

Arend Heyting developed an informal interpretation of the logical constants that tried to capture the principles of intuitionistic reasoning. This interpretation is known as the *Brouwer–Heyting–Kolmogorov–interpretation*. According to it, the logical constants are to be understood in the following way (Iemhoff, 2016):

- $\bot$ is not provable.
- A proof of $\phi \wedge \psi$ consists of a proof of ϕ and a proof of ψ.
- A proof of $\phi \vee \psi$ consists of a proof of A or a proof of ψ.
- A proof of $\phi \to \psi$ is a construction which transforms any proof of ϕ into a proof of ψ.
- A proof of $\exists x\phi(x)$ is given by presenting an element d of the domain and a proof of $\phi(d)$.
- A proof of $\forall x\phi(x)$ is a construction which transforms every proof that d belongs to the domain into a proof of $\phi(d)$.

Heyting also provided a formal logic that respects this interpretation, *intuitionistic logic* (we will here concentrate on the propositional part only).

It's actually quite simple to arrive at this system by modifying the system developed in Chapter 2 – all that's required is that we restrict sequents (and, via this, rules) to those that contain no more than one sentence on the right-hand side! For a few rules (explicitly, for WR, ER, CR and com) this means abandoning them altogether. For others it requires simply a restriction. For example, the negation rules must be restricted to the instances in which the right-hand side is empty in the bottom consecution; the cut rule to instances in which Δ is empty and Σ is a singleton, etc.

In the classical system, double negation elimination – i.e. the sequent $\neg\neg\phi \Rightarrow \phi$ – was easily derivable, e.g. as follows:

$$\dfrac{\dfrac{\phi \Rightarrow \phi}{\Rightarrow \neg\phi, \phi}\,\neg I_R}{\neg\neg\phi \Rightarrow \phi}\,\neg I_L$$

But the first step here is not intuitionistically acceptable. Of course, showing that *some* classically acceptable proof of $\neg\neg\phi \Rightarrow \phi$ fails intuitionistically (which is what we've just done) is a far cry from showing that there is *no* intuitionistically acceptable proof of $\neg\neg\phi \Rightarrow \phi$. To prove this stronger result, we would need to develop machinery beyond the scope of this text. But, as it turns out, this sequent is in fact not derivable at all in the single-consequence version of the system. (For proofs of this, see e.g. Dummett (2000).)

The new system *does*, however, allow us to derive *double negation introduction* – i.e. the sequent $\phi \Rightarrow \neg\neg\phi$ – for example as follows:

$$\dfrac{\dfrac{\phi \Rightarrow \phi}{\phi, \neg\phi \Rightarrow}\,\neg I_L}{\phi \Rightarrow \neg\neg\phi}\,\neg I_R$$

Notice also that while intuitionistic logic drops the *completeness* axiom $\Rightarrow \phi, \neg\phi$, it maintains the *consistency* axiom $\phi, \neg\phi \Rightarrow$.

The logic that results also blocks the proof of $\Rightarrow \phi \vee \neg\phi$. Classically, this was derivable as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\phi \Rightarrow \phi}{\phi \Rightarrow \phi \vee \neg\phi}\,R(\vee)}{\Rightarrow \neg\phi, \phi \vee \neg\phi}\,\neg I_R}{\Rightarrow \phi \vee \neg\phi, \phi \vee \neg\phi}\,R(\vee)}{\Rightarrow \phi \vee \neg\phi}\,CR$$

Again, the step involving $\neg I_R$ is not intuitionistically acceptable.

Since the proof rules for intuitionistic logic are just restricted forms of the proof rules for classical logic, it's clear that all intuitionistically derivable sequents are classically derivable. The converse, we've now seen, is not true.

Intuitionistic logic has some interesting properties. For example, Kurt Gödel proved the "Disjunction Property" for intuitionistic logic (IL):

$$\vdash_{\mathrm{IL}} \phi \vee \psi \text{ implies } \vdash_{\mathrm{IL}} \phi \text{ or } \vdash_{\mathrm{IL}} \psi$$

a principle which is violated in classical logic (since $(\phi \vee \neg\phi)$ also holds for instances of ϕ which aren't tautologies).

There are also several semantics developed for intuitionistic logic. The most prominent is perhaps Kripke's (1965). Some of these semantics (including Kripke's) are, however, only classical means to study intuitionistic logic, since the completeness proof for these logics is only provable by classical means (Kreisel, 1962). The situation is even a bit more curious than that. Not only is the completeness proof only classically available, completeness – for these semantics – also looks suspicious to the intuitionist, since by classical *and* intuitionistcally acceptable means it can be shown that if every intuitionistically valid formula is intuitionistically provable then a certain, intuitionistically highly implausible consequence follows (see Williamson, 2013).

As we said, intuitionistic logic was initially based on considerations in philosophy of mathematics, but it is a rich field in logic itself. Its relevance isn't confined to mathematical constructivism. As Michael Dummett argued, it should be a general insight to anti-realists that certain discourses will require logics that give up on the law of excluded middle (Dummett, 1991). In fact – for Dummett – figuring out which logical constants govern a discourse, and hence which logic is the right one for that discourse, was supposed to provide a way to settle the metaphysical dispute between realists and anti-realists bottom-up. This leads us to the question of what role logic should play in *philosophy*.

Logic and Philosophy

Being a philosophical subdiscipline, logic is of course of relevance to philosophy. Logic, as a theory, models notions that are of philosophical relevance, e.g. *logical truth*, *logical consequence*, *interpretation*, *inference*, *proof*, etc. It thereby sheds light on certain phenomena, such as the observation that some natural language sentences seem true just in virtue of their form, that some natural language arguments seem to be such that you can't rationally deny the conclusion if you accept the premises, which is not because of their particular content, but because of the form of these arguments, etc. – you know the story. Logic does that in a way in which many scientific theories explain real-life phenomena: it provides a partial model of the real-life phenomenon which, if adequate, displays the inner workings of the relevant mechanism that gives rise to the phenomenon, abstracting away from aspects of the real thing that are assumed not to be of any relevant influence. This is the "logic-as-modelling" view:

> The present claim is that a formal language is a mathematical model of a natural language, in roughly the same sense as, say, a Turing machine is a model of calculation, a collection of point masses is a model of a system of physical objects, and the Bohr construction is a model of an atom. In other words, a formal language displays certain features of natural languages, or idealizations thereof, while simplifying other features. (Shapiro, 2006, 49)

Consider another model. An architect scale model of a house uses materials such as paper, wood and clay, although the house it models is made of concrete, steel and glass, say. The architect's paper, wood, and clay are the equivalent of the logician's mathematical objects: sets, functions, ordered pairs, etc. The material that the architect is using in the model lacks many of the properties that the material of the house will have. But that's not a problem either since, again, the model does not need to represent all the aspects of the actual system modelled but only some of them. Also, a model does not need to represent the aspects that it *does* model by *exemplifying* them. This is why a tool of precision, such as mathematics, can be used to model a phenomenon of imprecision such as, for example, *vagueness* (see Cook, 2002; Shapiro, 2006).

Because models typically only model certain aspects of the actual phenomenon, while leaving out others, there might be several models of one and the same phenomenon, that are all "correct" or "adequate", simply because they are meant to model different aspects. Note that there doesn't have to be one supermodel that models all aspects and is thus, in some sense, more "correct" than all others. It is often the very point of modelling to leave out certain details in order to represent those aspects of the actual phenomenon that matter for understanding what one is interested in. In these cases any more "complete" model wouldn't be better, it would serve its function (as a model of that aspect) worse than the more limited model (see Cook, 2010). We explained in Chapter 7 how this "logic-as-modelling" view leads to a certain kind of pluralism about logic.

But the role of logic in philosophy is far wider than merely providing a theory of such concepts as *logical consequence*. Many philosophical texts contain symbols from formal logic, even though in these texts the formulas are not part of a model of any natural language (or other phenomenon). Instead they are part of the *technically enriched* language in which the philosophical text is written. This might be useful, if one wants to make, for example, complicated quantified claims and tries to avoid scope ambiguities, or wants to make explicit what one is quantifying over, etc. In these cases, logical notation functions like abbreviations, or as explicitly defined theoretical terms.

As we have seen above, logic as a theory also plays a role in other empirical theories which are of interest to philosophers. Formal semantics is of interest to philosophers of language; theories of reasoning, as they are developed in cognitive science, are of interest to epistemologists. In fact, using logic to model *epistemic processes* is not uncommon at all in philosophy. Analytic epistemology is traditionally interested in *rational reconstructions* of cognitive processes. The purpose of these rational reconstructions is to make plain how a certain cognitive process might eventually result in knowledge or justified beliefs, etc., if we pre-theoretically think that we have such knowledge or such justified beliefs. Typically a rational reconstruction assumes some (more or less) unproblematic basis of knowledge and some justification-preserving inference pattern and then goes on to show how these two suffice to generate the *explicandum*.

The role of these justification-preserving inference patterns seems crucial. It is not enough just to know that so far we have been quite successful in reasoning from basis *X* with pattern *Y*; the philosophical analysis should tell us *why* that is so. This explanatory function is usually satisfied by delimiting the choice of inference patterns (based on *a priori* considerations).

In, for example, *modal epistemology* we try to apply the project of analytic epistemology to our knowledge of necessities and possibilities. How do we know what is (merely) possible? Mere possibilities can't be directly observed, and neither can they be derived from our knowledge of what's actual.[4] Likewise, what is necessary goes beyond what is actually the case. But then, how can we know of any (nontrivial) necessities?

Consider as an example Timothy Williamson's account. Williamson (2007) observes that we seem to have a capacity to evaluate *counterfactuals*. We can even explain why we should have such a capacity and why we should consider it to be reliable (on the basis of the role that it plays in our decision making). We often consider *what would be the case if* and then base our decisions (largely successfully) on the projected outcomes. Williamson then goes on to show that modal logic (of the strength of S5) is a subsystem of counterfactual logic. In other words, knowing the truth of certain counterfactuals is sufficient for knowing corresponding possibilities and necessities.

In order for such an account to *explain* our problematic knowledge more is needed than merely showing that there is *some* logical route from an unproblematic kind of knowledge to the kind of knowledge we want to explain. It also needs to be shown that this logical route is psychologically real. In our example, this would require showing that our knowledge of

[4] Mere possibilities can't even be straightforwardly derived from knowledge of what is necessary, if the language in which we reason has at least the expressive powers of polyadic first-order logic (see Cohnitz, 2012).

necessities and possibilities is actually arrived at by reasoning through certain counterfactuals (Jenkins, 2008; Cohnitz, 2012).

The most prominent application of logic in philosophy is, however, considered to lie in the *formal reconstruction of arguments*. At least most introductory textbooks to philosophical logic "sell" logic that way. Logic is then seen as a central *methodological tool* of philosophical analysis. We can check whether arguments (our own as well as those of other philosophers) are logically valid, find points of criticism if they are not, discover hidden premises, etc.

However, it quickly becomes apparent to students in such introductory logic courses that matters aren't quite so straightforward. First of all, even though the "gold standard" for good arguments might be that of deductive validity, philosophy often doesn't conform to that standard. Philosophical argumentation is more often *abductive* rather than *deductive*. We don't possess a set of shared, general principles from which we can simply deduce the solutions to philosophical problems. Often philosophers have to go with what's *plausible* given certain other assumptions, rather than with what follows from such assumptions with necessity. Still, such arguments can be good arguments; as we already said, they might be the best arguments available in a certain domain. Hence, *being deductively valid* is not actually a necessary condition for *being a good argument* in philosophy, and thus showing by formal means that an argument is not deductively valid is a bit less exciting than it might initially have seemed.

But even if deductive validity was a necessary condition for goodness of arguments, formal logic is only a very limited tool for establishing invalidity. At most one can show that an argument is invalid *on that level of logical analysis*, leaving it open that there might be a deeper level of logical analysis on which the argument is deductively valid after all. For example, showing that an argument is invalid if reconstructed in propositional logic leaves open whether the argument might be deductively valid in first-order predicate logic. Showing that an argument is invalid in first-order predicate logic leaves open whether it might be valid in second-order logic, or modal logic, etc.[5]

As should be clear, being formally (and thus deductively) valid is not *sufficient* for being a good argument either. Arguments that are deductively

[5] You remember this discussion from Chapter 4, the Debunker's view on logical form.

valid also have to have true premises in order to be *sound*, and might have to satisfy further constraints in order to be convincing, etc. Rather, being deductively valid is something like an *INUS-condition* for being a good argument, an **i**nsufficient, but **n**ecessary part of an **u**nnecessary but **s**ufficient condition.

So, logic has its limits as a general methodological tool. Still, having a general methodology in philosophy, even if it is of only limited applicability, might seem highly desirable. Philosophers challenge everything; no assumption, no basic belief is safe from being scrutinized. Having at least logic as a framework in which philosophy proceeds (to some degree) might be seen as the only thing that keeps philosophical enquiry from ending up in intellectual anarchy and chaos. Can logic really play the role of such a framework for philosophy? What constraints does that put on logic?

In Chapter 5 we discussed the view that logic should be metaphysically neutral. If we want logic to provide something like a general framework in which philosophy can proceed, then that framework should be a completely *neutral arbiter*: it shouldn't have any metaphysical or other content which could preclude a fair evaluation of any philosophical view. As we have seen, logic is not like that. Virtually every logical principle has been challenged at some point and the basis of these challenges have often been metaphysical considerations. Thus, logic isn't completely neutral in that sense. Every logical framework makes some metaphysical assumptions. Fortunately that doesn't mean that logic must be useless for philosophy. We have also seen that there is a wide variety of available logical frameworks at the level of logical theory. Thus, if one is worried about the metaphysical presuppositions of a logic being in the way of its neutrality given a certain topic, one might chose a different framework which doesn't have the problematic presupposition.

There are further properties that might be desirable in a logic, if we take the application of logic in philosophy as our primary motivation. For example, philosophical theories – metaphysical theories in particular – are, arguably, intended as unrestricted claims. The physicalist claim that *everything* is physical is not supposed to apply only to a suitably restricted domain; it is meant to talk about *absolutely everything*. If that's so, then the logic of philosophy should allow for unrestricted quantification (cf. our discussion in Chapter 2).

Likewise, when philosophers of language want to talk about truth, they don't just want to talk about *truth-in-L*, but about *truth* in general, as it concerns all languages. If this is not possible within standard logic (because of the Liar Paradox), then the logic of philosophy perhaps needs to be non-classical:

> Especially philosophy cannot restrict itself to non-universal languages. Philosophy does not want to deal only with the structure or conditions of talking in some specific language or languages of some kind, but aims at a theory of the basic structures and conditions of having a language *in general*. This requires the corresponding resources to express the universal claims. Universal theories of meaning, truth, knowledge etc. [are not possible] if we can talk only from some meta-language "down" to some distinct object-language. [. . .] Our concept of language, therefore, involves unity and universality. There has to be a set of properties defining what a language is. These properties are preserved in change or translation. Without semantic closure we would not be *able* to elucidate a concept that we seem to have! So I take it that we need semantic closure. Nothing, but dialetheism seems to be able to deliver it. (Bremer, 2008, 212–213)

These are at least possible considerations that one could raise for the evaluation of a particular logic as a methodological framework for philosophy.

Logic as a Science

Logic is special; logic is general, and basic, it has normative force and is certain. There are also other things that have such properties, but logic is special in possessing these properties in the absolute extreme. As Gila Sher (1999) argues, this makes it hard to give an explanatory account of logic and its properties. Let's begin with logic's generality. Logic supposedly applies to every subject matter. But if logic applies to everything conceivable, then it has no subject matter of its own, and thus doesn't say anything in particular about anything. Hence logic is apparently empty. Furthermore, if logic is indeed universally applicable, then there is no place outside logic, no "vantage point from which to explain logic" (Sher, 1999, 210).

Logic's basicness leads to similar puzzles. Logic is the most fundamental theory. There is no knowledge possible without logic. But if logic and its concepts are the most basic and fundamental, then there is nothing from which logic itself could be explained, no concepts “further down” in the hierarchy that could ground or explain the concepts or properties of logic.

Systems of norms other than logic are all constrained by the norms of logic. In contrast, the norms of logic are not constrained by any other system of norms. But how could the authority of the norms of logic be explained if not by showing how it flows from other, more basic norms?

Likewise, logic is absolutely certain. We can’t point to anything more certain from which we could begin to explain the certainty of logic. Logic seems to be a unique thing, quite peculiar among the sciences.

Logic is not special. In the same paper in which Gila Sher lists these apparent obstacles for any explanation of logic, she also provides an account of logic as a theory of the formal that would explain all these features of logic, including the excessiveness in which logic seems to possess them. As we have seen in Chapter 4, Sher demarcates the formal via an *invariance principle*. “Formal properties do not distinguish between isomorphic structures of objects” (Sher, 1999, 233). If that’s so, then there can’t be a domain of objects for which the laws of logic do not hold (which explains logic’s generality). Since logic is so general, it holds in all domains, including the domains of all other sciences, but since formal properties do not distinguish the properties of other sciences, logic is independent of these (which explains its basicness and *mutatis mutandis* also its superior normative force). Thanks to the invariance principle, logic seems in particular to be invariant with respect to empirical differences between structures, which explains logic’s high degree of epistemic stability.

Perhaps logic is even less special. Perhaps the difference between logic and other sciences is more a matter of degree. As Penelope Maddy (2014) shows, we could have arrived at a theory explaining the reliability of logical inferences in situations that obey certain formal constraints without assuming any particularly “philosophical” machinery (such as notions of possible worlds, or an account of *a priori* knowledge, etc.), as a “perfectly ordinary answer to a perfectly ordinary question”. Her unassuming enquirer, the so-called “Second Philosopher”, begins with a simple question:

> What’s hidden in my hand is either an ordinary dime or a foreign coin of a type I’ve never seen. (I drew it blindfolded from a bin filled with just these

> two types of objects.) It's not a dime. (I can tell by the feel of it.) Then, obviously, it *must* be a foreign coin! But what makes this so?
> (Maddy, 2014, 93)

The Second Philosopher quickly observes that the reliability of such inferences is independent of all but the most general structural features of the situation, notices how conjunctions and disjunctions interrelate and that properties can hold for some objects or sometimes for all. Via a three-valued logic, which the Second Philosopher calls "rudimentary logic", she eventually arrives at a system formally like classical logic which works as a widely applicable idealization to all situations that instantiate formal structure. From developmental psychology she learns why inferences in accordance with that system would strike us as being obvious: we develop early on a capacity to detect formal structure in the world, a capacity we have thanks to the fact that we typically interact with aspects of the world instantiating formal structure.

> [The Second Philosopher's] answer doesn't deliver on the usual philosophical expectations: the reliability of the inference is contingent, our knowledge of it is only minimally *a priori* at best. The account itself results from plain empirical enquiry, which may lead some to insist that it isn't philosophy at all. (Maddy, 2014, 108)

Views like this, which hold that logic is continuous with science in content and method, have come to be called "anti-exceptionalist" (Hjortland, 2017). According to such views, not only is philosophy an abductive discipline (as we have already argued above), but even logic follows the abductive methodology of the sciences. Logics are revisable and can be compared which each other as to how well they fit the evidence, but also with respect to other theoretical virtues, such as "strength, simplicity, elegance, and unifying power" (Williamson, 2017, 14). We have encountered such anti-exceptionalist views already in Chapters 5 and 6.

According to some anti-exceptionalists, logic doesn't have a special normative status either, nor does it have a special content (thus logic is neither describing the psychology of deductive reasoning, nor in any sense about language). Williamson, for example, holds a *deflationary* view according to which a logical theory consists of sentences that are unrestricted universal generalizations (in fact, the universal closures of valid arguments). These sentences are about neither language, concepts nor reasoning, but about the world (Williamson, 2017).[6]

Like Maddy, Williamson holds that – even on such a revisionist view about the nature of logical theory and methodology – classical logic should come out as the comparatively best theory. He argues that the best case for alternatives to classical logic can be made on the basis of the semantic paradoxes. Classical logic and unrestricted disquotation are inconsistent – we can't have them both. In Williamson's abductive methodology, the question now becomes which theoretical virtues on either side can help us in deciding what we want to give up. Both unrestricted classical logic and unrestricted disquotation are desirable.

Eventually, Williamson decides to favour the more fundamental theory (as a general methodological principle of the sciences) and argues that classical logic is more fundamental than disquotation, since the former is integral to mathematics, which is integral to all our best scientific theories, such that a revision of classical logic would lead to major changes to theories in all sciences. Disquotation, on the other hand, is only concerned with language and thus a lot less fundamental and easier to revise (see Hjortland (2017) for a reconstruction of Williamson's argument).

This leads us back to a question we encountered at the very beginning of this chapter. Is logic really fundamental for the sciences, and, if so, then in what sense?

Of course, logical reasoning *is* fundamental to all sciences. Plausibly, classical logic is in that sense indeed an integral part of mathematics. But that doesn't make *classical logical theory* an integral part of mathematics. As Hjortland (2017) argues, mathematics was already done before classical logic was properly captured in a formal theory and also nowadays largely proceeds without explicitly citing logical theory. The fact that mathematical reasoning often instantiates classical principles is compatible with classical logic being restricted to the mathematical domain (and hence does not support the kind of unrestricted universal generalizations that Williamson considers logic to consist of).

6 There are also *non-deflationary* anti-exceptionalists, such as Priest (2016) who would still hold that logical theories should be about the familiar logical notions, such as validity, logical consequence, consistency, etc. (Hjortland, 2017).

But if logic is not the most fundamental theory for that reason, then this is just another sense in which logic isn't special as a theory.

Questions

1. We said that some semanticists invoke the Principle of Compositionality in order to explain how finite human beings can understand languages with infinitely many expressions with different meanings. How is this explanation achieved? Which feature of compositionality is required for such an explanation?
2. Assume that a "Panglossian" approach would succeed and it could be shown that human beings reason by and large correctly. Would that have implications for psychologism (as a metaphysical view on logic)?
3. In the discussion of intuitionistic proof theory we suggested that it is difficult to show that the provability of a certain classical theorem is blocked. What would be needed to show that a certain formula is not provable in such an intuitionistic system?

Bibliography

Azzouni, Jody, 2014. A defense of logical conventionalism. In Penelope Rush (ed.), *The Metaphysics of Logic*, pp. 32–48. Cambridge: Cambridge University Press.

Barwise, Jon, 1997. Information and impossibilities. *Notre Dame Journal of Formal Logic*, 38: 488–515.

Bealer, George, 2000. A priori knowledge. In Jaakko Hintikka, Robert Neville, Ernest Sosa and Alan Olson (eds), *The Proceedings of the Twentieth World Congress of Philosophy*, vol. 5, pp. 1–12. Charlottesville, VA: Philosophy Documentation Center.

Beall, Jc, 2009. *Spandrels of Truth*. Oxford: Oxford University Press.

Beall, Jc, 2011. Multiple-conclusion LP and default classicality. *Review of Symbolic Logic*, 4: 326–336.

Beall, Jc, 2013. LP^+, $K3^+$, FDE^+, and their 'classical collapse'. *Review of Symbolic Logic*, 6 (4): 742–754.

Beall, Jc and Logan, Shay, 2017. *Logic: The Basics*, 2nd edn. Routledge.

Beall, Jc and Murzi, Julien, 2013. Two flavors of Curry's paradox. *Journal of Philosophy*, 110 (3): 143–165.

Beall, Jc and Restall, Greg, 2000. Logical pluralism. *Australasian Journal of Philosophy*, 78: 475–493.

Beall, Jc and Restall, Greg, 2001. Defending logical pluralism. In John Woods and Bryson Brown (eds), *Logical Consequence: Rival Approaches*, pp. 1–22. Oxford: Hermes Science Publishing.

Beall, Jc and Restall, Greg, 2006. *Logical Pluralism*. Oxford: Oxford University Press.

Beall, Jc, Forster, Thomas and Seligman, Jeremy, 2013. A note on freedom from detachment in the logic of paradox. *Notre Dame Journal of Formal Logic*, 54: 15–20.

Belnap, Nuel D., 1962. Tonk, plonk and plink. *Analysis*, 23: 130–134.

Belnap, Nuel D., 1977. A useful four-valued logic. In J. Michael Dunn and George Epstein (eds), *Modern Uses of Multiple-Valued Logic*, pp. 5–37. Dordrecht: Springer.

Béziau, Jean-Yves, 1995. Universal logic. In T. Childers and O. Majer (eds), *Logica'94: Proceedings of the 8th International Symposium*, pp. 73–93. Montreal: Filosofia.

Béziau, Jean-Yves, 2001. From paraconsistent logic to universal logic. *Sorites*, 12: 5–32.

Béziau, Jean-Yves, 2006. 13 questions about universal logic. *Bulletin of the Section of Logic*, 35 (2–3): 133–150.

Béziau, Jean-Yves, 2010. What is a logic? –Towards axiomatic emptiness. *Logical Investigations*, 16: 272–279.

Boghossian, Paul, 2000. Knowledge of logic. In Paul Boghossian and Christopher Peacocke (eds), *New Essays on the Apriori*, pp. 229–254. Oxford: Oxford University Press.

Bonnay, Denis, 2014. Logical constants, or how to use invariance in order to complete the explication of logical consequence. *Philosophy Compass*, 9: 54–65.

Boolos, George, 1975. On second-order logic. *Journal of Philosophy*, 72: 509–527.

Boolos, George, 1984. To be is to be a value of a variable (or to be some values of some variables). *Journal of Philosophy*, 81: 430–449.

Bremer, Manuel, 2003. Do logical truths carry information? *Minds and Machines*, 13: 567–575.

Bremer, Manuel, 2008. Why and how to be a dialetheist. *Studia Philosophica Estonica*, 1(2): 208–227.

Bremer, Manuel, 2014. Restall and Beall on logical pluralism: A critique. *Erkenntnis*, 79: 293–299.

Bremer, Manuel and Cohnitz, Daniel, 2004. *Information and Information Flow: An Introduction*. Berlin: De Gruyter.

Broome, John, 2000. Normative requirements. In Jonathan Dancy (ed.), *Normativity*, pp. 78–99. Oxford: Blackwell.

Brouwer, L. E. J., 1981. *Brouwer's Cambridge Lectures on Intuitionism*. Cambridge: Cambridge University Press.

Bueno, Otávio, 2001. Can a paraconsistent theorist be a logical monist? In Walter A. Carnielli, Marcelo Coniglio and Ítala M.L D'Ottaviano (eds), *Paraconsistency: The Logical Way to the Inconsistent*, pp. 535–552. New York: Marcel Dekker.

Bueno, Otávio and Shalkowski, Scott A., 2009. Modalism and logical pluralism. *Mind*, 118 (470): 295–321.

Burge, Tyler, 2007. On knowledge and convention. In *Foundations of Mind*, pp. 32–37. Oxford: Oxford University Press.

Carnap, Rudolf, 1937. *The Logical Syntax of Language*. London: Routledge & Kegan Paul.

Carnap, Rudolf, 1956. Empiricism, semantics, and ontology. In *Meaning and Necessity*, 2nd edn, pp. 203–221. Chicago, IL: University of Chicago Press.

Carnap, Rudolf and Bar-Hillel, Yehoshua, 1952. An outline of a theory of semantic information. Technical Report 247, MIT Research Labaratory of Electronics.

Carroll, Lewis, 1895. What the tortoise said to Achilles. *Mind*, 4: 278–280.

Chalmers, David, 2015. Why isn't there more progress in philosophy? *Philosophy*, 90: 3–31.

Coffa, Alberto, 1986. From geometry to tolerance: Sources of conventionalism in nineteenth-century geometry. In Robert G. Colodny (ed.), *From Quarks to Quasars: Philosophical Problems of Modern Physics*, pp. 7–3. Pittsburgh, PA: University of Pittsburgh Press.

Cohnitz, Daniel, 2005. Is compositionality an apriori principle? In Markus Werning, Edouard Machery and Gerhard Schurz (eds), *The Compositionality of Concepts and Meanings: Foundational Issues*, pp. 23–58. Berlin: De Gruyter.

Cohnitz, Daniel, 2012. The logic(s) of modal knowledge. In Greg Restall and Gillian Russell (eds), *New Waves in Philosophical Logic*, pp. 63–83. Basingstoke: Palgrave Macmillan.

Cohnitz, Daniel and Haukioja, Jussi, 2015. Intuitions in philosophical semantics. *Erkenntnis*, 80: 617–641.

Cohnitz, Daniel and Rossberg, Marcus, 2006. *Nelson Goodman*. London: Routledge.

Collins, John, 2009. The limits of conceivability: logical cognitivism and the language faculty. *Synthese*, 171: 175–194.

Colyvan, Mark, 2006. Naturalism and the paradox of revisability. *Pacific Philosophical Quarterly*, 87(1):1–11.

Cook, Roy, 2002. Vagueness and mathematical precision. *Mind*, 111(442):227–247.

Cook, Roy, 2004. What's wrong with tonk (?). *Journal of Philosphical Logic*, 34: 217–226.

Cook, Roy, 2010. Let a thousand flowers bloom: A tour of logical pluralism. *Philosophy Compass*, 5: 492–504.

Cook, Roy, 2014. There is no paradox of logical validity. *Logica Universalis*, 8 (3-4): 447–467.

Corry, Leo, 2004. *Modern Algebra and the Rise of Mathematical Structures*, 2nd edn. Basel: Birkhäuser.

Cosmides, L., 1989. The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31: 187–276.

Cosmides, Leda and Tooby, John, 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.

Cresswell, M., 1973. *Logics and Languages*. London: Methuen.

da Costa, Newton C. A., 1974. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15: 497–510.

da Costa, Newton C. A., 1979. *Ensaio sobre os fundamentos da Lógica*. São Paulo: Editora HUCITEC.

da Costa, Newton C. A., 1997. *Logiques classiques et non classiques: Essai sur les fondements de la logique*. Paris: Masson.

Došen, Kosta, 1989. Logical constants as punctuation marks. *Notre Dame Journal of Formal Logic*, 30: 362–381.

Dummett, Michael, 1978. The justification of deduction. In *Truth and Other Enigmas*, pp. 166–185. Cambridge, MA: Harvard University Press.

Dummett, Michael, 1991. *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.

Dummett, Michael A. E., 1998. Truth from the constructive standpoint. *Theoria*, 64: 122–138.

Dummett, Michael A. E., 2000. *Elements of Intuitionism*. Oxford: Oxford University Press.

Dutilh Novaes, Catarina, 2015. A dialogical, multi-agent account of the normativity of logic. *Dialectica*, 69: 587–609.

Ebert, Philip A. and Rossberg, Marcus, 2007. What is the purpose of neologicism? *Travaux de Logique*, 18:33–61.

Estrada-González, Luis, 2011. On the meaning of connectives (apropos of a non-necessitarianist challenge). *Logica Universalis*, 5: 115–126.

Estrada-González, Luis, 2012. Models of possibilism and trivialism. *Logic and Logical Philosophy*, 21: 175–205.

Etchemendy, John, 1988. Tarski on truth and logical consequence. *Journal of Symbolic Logic*, 53: 51–79.

Field, Hartry, 1991. Metalogic and modality. *Philosophical Studies*, 62: 1–22.

Field, Hartry, 1996. The a prioricity of logic. *Proceedings of the Aristotelian Society, New Series*, 96: 359–379.

Field, Hartry, 1998. Epistemological nonfactualism and the a prioricity of logic. *Philosophical Studies*, 92: 1–24.

Field, Hartry, 2004. The revision of logical laws. unpublished. Available at: https://pdfs.semanticscholar.org

Field, Hartry, 2008. *Saving Truth from Paradox*. Oxford: Oxford University Press.

Field, Hartry, 2009a. What is the normative role of logic? *Proceedings of the Aristotelian Society, Supplementary Volume*, 83(1):251–268.

Field, Hartry, 2009b. Pluralism in logic. *Review of Symbolic Logic*, 2: 342–359.

Fjellstad, Andreas, 2015. How a semantics for tonk should be. *Review of Symbolic Logic*, 8: 488–505.

Frankowski, Szymon, 2004a. Formalization of a plausible inference. *Bulletin of the Section of Logic*, 33: 41–52.

Frankowski, Szymon, 2004b. p-consequence versus q-consequence operations. *Bulletin of the Section of Logic*, 33: 197–207.

Frapolli, Maria J., 2012. ¿Qué son las constantes lógicas? *Critica*, 44: 65–99.

Frege, Gottlob, 1961 [1884]. *Die Grundlagen der Arithmetik: eine logisch mathematische Untersuchung über den Begriff der Zahl*. Hildesheim: Georg Olms.

Frege, Gottlob, 1962 [1893]. *Grundgesetze der Arithmetik*, volume I. Hildesheim: Georg Olms Verlags.

Frege, Gottlob, 1980. *The Foundations of Arithmetic*, 2nd edn. Evanston, IL: Northwestern University Press.

French, Rohan, 2016. Structural reflexivity and the paradoxes of self-reference. *Ergo*, 2(13):299–328.

Geach, P. T., 1965. Assertion. *Philosophical Review*, 74: 449–465.

Gentzen, Gerhard, 1969. *The Collected Papers of Gerhard Gentzen*, ed. M. E. Szabo, Studies in Logic and Foundations of Mathematics Vol. 55. Amsterdam: North Holland.

Gigerenzer, G., 2000. *Adaptive Thinking: Rationality and the Real World*. Oxford: Oxford University Press.

Goldman, Alvin, 2007. Philosophical intuitions, their target, their source, and their epistemic status. *Grazer Philosophische Studien*, 74: 1–26.

Gomez-Torrente, Mario, 2002. The problem of logical constants. *Bulletin of Symbolic Logic*, 8: 1–37.

Goodman, Nelson, 1955. *Fact, Fiction, Forecast*. Cambridge, MA: Harvard University Press.

Gottlieb, Paula, 2015. Aristotle on non-contradiction. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, summer 2015 edition. Available at: https://plato.stanford.edu/archives/sum2015/entries/aristotle-noncontradiction.

Haack, Susan, 1978. *Philosophy of Logics*. Cambridge: Cambridge University Press.

Haack, Susan, 1996. *Deviant Logic, Fuzzy Logic: Beyond the Formalism*. Chicago, IL: University of Chicago Press.

Hacking, Ian, 1979. What is logic? *Journal of Philosophy*, 76: 285–319.

Halbach, Volker, 2010. *The Logic Manual*. Oxford: Oxford University Press.

Halbach, Volker, 2011. *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Hanna, Robert, 2006a. *Rationality and Logic*. Cambridge, MA: MIT Press.

Hanna, Robert, 2006b. Rationality and the ethics of logic. *Journal of Philosophy*, 103: 67–100.

Hardegree, Gary M., 2005. Completeness and super-valuations. *Journal of Philosophical Logic*, 34: 81–95.

Harman, Gilbert, 1984. Logic and reasoning. *Synthese*, 60: 107–127.

Harman, Gilbert, 1986. *Change in View*. Cambridge, MA: MIT Press.

Hintikka, Jaakko, 1970. Surface information and depth information. In Jaakko Hintikka and P. Suppes (eds), *Information and Inference*, pp. 263–297. Dordrecht: Reidel.

Hjortland, Ole, 2016. Varieties of logic. *Philosophical Quarterly*, 66: 646–648.

Hjortland, Ole, 2017. Anti-exceptionalism about logic. *Philosophical Studies*, 174(3): 631–658.

Hodes, Harold, 2004. On the sense and reference of a logical constant. *Philosophical Quarterly*, 54: 134–165.

Hughes, G. E. and Cresswell, M. J., 1996. *A New Introduction to Modal Logic*. London: Routledge.

Iemhoff, Rosalie, 2016. Intuitionism in the philosophy of mathematics. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2016 edition. Available at: https://plato.stanford.edu/archives/win2016/entries/intuitionism.

Jago, Mark, 2013. The content of deduction. *Journal of Philosphical Logic*, 42: 317–334.

Jeffrey, Richard, 1967. *Formal Logic: Its Scope and Limits*. New York: McGraw Hill.

Jenkins, C. S. I., 2008. Modal knowledge, counterfactual knowledge and the role of experience. *The Philosophical Quarterly*, 58: 693–701.

Johnson-Laird, P. N., Legrenzi, P. and Legrenzi, M. S, 1972. Reasoning and a sense of reality. *British Journal of Psychology*, 63: 395–400.

Katz, Jerrold J., 1998. *Realistic Rationalism*. Cambridge, MA: MIT Press.

Ketland, Jeffrey, 2012. Validity as a primitive. *Analysis*, 72 (3): 421–430.

Kleene, S. C., 1952. *Introduction to Metamathematics*. New York: Van Nostrand.

Kreisel, Georg, 1962. On weak completeness of intuitionistic predicate logic. *Journal of Symbolic Logic*, 27:139–158.

Kreisel, Georg, 1972. Informal rigour and completness proofs. In Imre Lakatos (ed.), *Problems in the Philosophy of Mathematics*, pp. 138–171. Amsterdam: North-Holland.

Kripke, Saul, 1965. Semantical analysis of intuitionistic logic. In J. Crossley and M. A. E. Dummett (eds), *Formal Systems and Recursive Functions*, pp. 92–130. Amsterdam: North-Holland.

Künne, Wolfgang, 2003. *Conceptions of Truth*. Oxford: Oxford University Press.

Kusch, Martin, 2015. Psychologism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2015 edition. Available at: https://plato.stanford.edu/archives/win2015/entries/psychologism.

Lakatos, Imre, 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: Cambridge University Press.

Lapointe, Sandra, 2014. Bolzano's logical realism. In Penelope Rush (ed.), *The Metaphysics of Logic*, pp. 189–208. Cambridge: Cambridge University Press.

Lewis, David, 1969. *Convention: A Philosophical Study*. Cambridge: Cambridge University Press.

Lewis, David, 1984. Putnam's paradox. *Australasian Journal of Philosophy*, 62: 221–236.

Linnebo, Øystein, 2014. Plural quantification. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2014 edition. Available at: https://plato.stanford.edu/archives/fall2014/entries/plural-quant/.

Löbner, Sebastian, 2013. *Understanding Semantics*, 2nd edn. London: Routledge.

MacFarlane, John, 2000. What does it mean to say that logic is formal? PhD thesis, University of Pittsburgh.

MacFarlane, John, 2004. In what sense (if any) is logic normative for thought? unpublished manuscript, available online.

MacFarlane, John, 2015. Logical constants. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2015 edition. Available at: https://plato.stanford.edu/entries/logical-constants/.

Maddy, Penelope, 2014. A second philosophy of logic. In Penelope Rush (ed.), *The Metaphysics of Logic*, pp. 93–108. Cambridge: Cambridge University Press.

Malinowski, Grzegorz, 1990a. Q-consequence operation. *Reports on Mathematical Logic*, 24: 49–59.

Malinowski, Grzegorz, 1990b. Towards the concept of logical many-valuedness. *Folia Philosophica*, 7: 97–103.

Mates, Benson, 1952. Synonymity. In Leonard Linsky (ed.), *Semantics and the Philosophy of Language*, pp. 111–38. Champaign, IL: University of Illinois Press.

Mates, Benson, 1965. *Elementary Logic*. Oxford: Oxford University Press.

Maudlin, Tim, 2004. *Truth and Paradox: Solving the Riddles*. Oxford: Oxford University Press.

McCarthy, T., 1987. Modality, invariance, and logical truth. *Journal of Philosophical Logic*, 16: 423–443.

McGee, V., 1996. Logical operations. *Journal of Philosophical Logic*, 25: 567–580.

Mill, J. S., 1843. *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. London: John W. Parker.

Miller, Alexander, 2007. *Philosophy of Language*, 2nd edn. London: Routledge.

Millikan, R. G., 1998. Language conventions made simple. *Journal of Philosophy*, 95: 61–180.

Mortensen, C., 1989. Anything is possible. *Erkenntnis*, 30: 319–337.

Mortensen, Chris, 2005. It isn't so, but could it be? *Logique et Analyse*, 48: 189–192.

Morton, A., 1973. Denying the doctrine and changing the subject. *Journal of Philosophy*, 70: 503–510.

Moschovakis, Joan, 2015. Intuitionistic logic. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2015 edition. Available at: https://plato.stanford.edu/archives/spr2015/entries/logic-intuitionistic.

Murzi, Julien, 2014. The inexpressibility of validity. *Analysis*, 74: 65–81.

Neurath, Otto, 1973. Anti-Spengler: Dedicated to the young and the future they shape. In Marie Neurath and Robert S. Cohen (eds), *Empiricism and Sociology*, pp. 158–213. Dordrecht: Reidel.

Padro, Romina, 2015. What the tortoise said to Kripke: the adoption problem and the epistemology of logic. PhD thesis, City University of New York.

Paoli, Francesco, 2003. Quine and Slater on paraconsistency and deviance. *Journal of Philosophical Logic*, 32 (5): 531–548.

Paoli, Francesco, 2007. Implicational paradoxes and the meaning of logical constants. *Australasian Journal of Philosophy*, 85 (4): 553–579.

Paseau, Alexander, 2007. Review: *Logical Pluralism. Mind*, 116 (462): 391–396.

Pelletier, Francis Jeffry, 1994. The principle of sematic compositionality. *Topoi*, 13: 11–24.

Pelletier, Francis Jeffry, Elio, Renée and Hanson, Philip, 2008. Is logic all in our heads? From naturalism to psychologism. *Studia Logica*, 86: 1–65.

Petersen, Uwe, 2000. Logic without contraction as based on inclusion and unrestricted abstraction. *Studia Logica*, 64 (3): 365–403.

Philipse, Herman, 1989. Psychologism and the prescriptive function of logic. In M. A. Notturno (ed.), *Perspectives on Psychologism*, pp. 58–73. Leiden: E. J. Brill.

Popper, Karl, 1947. New foundations for logic. *Mind*, 56: 193–235.

Prawitz, Dag, 1965. *Natural Deduction: A Proof-theoretical Study*. Stockholm: Almqvist and Wiksell.

Prawitz, Dag, 1982. Beweise und die Bedeutung und Vollständigkeit der logischen Konstanten. *Conceptus*, 16: 31–44.

Priest, Graham, 1979. The logic of paradox. *Journal of Philosophical Logic*, 8 (1): 219–241.

Priest, Graham, 2003. On alternative geometries, arithmetics, and logics; a tribute to Łukasiewicz. *Studia Logica*, 74: 441–468. Reprinted with some modifications in Priest (2006), ch. 10; quotes are from this edition.

Priest, Graham, 2006. *Doubt Truth to Be a Liar*. Oxford: Oxford University Press.

Priest, Graham, 2008. *An Introduction to Non-Classical Logic*, 2nd edn. New York: Cambridge University Press.

Priest, Graham, 2014. Revising logic. In P. Rush (ed.), *The Metaphysics of Logic*, pp. 211–223. Cambridge: Cambridge University Press.

Priest, Graham, 2016. Logical disputes and the a priori. *Logique and Analyse*, 59(236):347–366.

Priest, Graham and Routley, Richard, 1982. Lessons from Pseudo Scotus. *Philosophical Studies*, 42: 189–199.

Prior, A., 1960. The runabout inference-ticket. *Analysis*, 21: 38–39.

Prior, A. N., 1971. Platonism and quantification. In A. N. Prior, P. T. Geach and J. P. Kenny (eds), *Objects of Thought*. Oxford: Oxford University Press.

Putnam, Hilary, 1957. Three-valued logic. *Philosophical Studies*, 8 (5). Reprinted in Putnam (1979, 166–173).

Putnam, Hilary, 1962. The analytic and the synthetic. In Herbert Feigl and Grover Maxwell (eds), *Minnesota Studies in the Philosophy of Science*, pp. 358–397. Minneapolis, MN: University of Minnesota Press. Reprinted in *Mind, Language, and Reality: Philosophical Papers Volume 2*, New York: Cambridge University Press, 1975, pp. 33–69; quotes are from this edition.

Putnam, Hilary, 1968. Is logic empirical? In Robert S. Cohen and Marx W. Wartofsky (eds), *Boston Studies in the Philosophy of Science*, volume 5. Dordrecth: Reidel. Reprinted in Putnam (1979, 216–241) as "The logic of quantum mechanics"; quotes are from this edition.

Putnam, Hilary, 1975. The logic of quantum mechanics. In *Mathematics, Matter and Method: Philosophical Papers Volume 1*, pp. 174–197. Cambridge: Cambridge University Press.

Putnam, Hilary, 1979. *Mathematics, Matter, and Method*, 2nd edn. Cambridge: Cambridge University Press.

Putnam, Hilary, 1980. Models and reality. *Journal of Symbolic Logic*, 45: 421–444.

Quine, W. V. O., 1953a. Two dogmas of empiricism. In *From a Logical Point of View*, pp. 20–46. Cambridge, MA: Harvard University Press.

Quine, W. V. O., 1953b. Three grades of modal involvement. In *Proceedings of the XIth International Congress of Philosophy*, vol. 14. Amsterdam: North-Holland.

Quine, W. V. O., 1960. *Word and Object*. New York: Wiley.

Quine, W. V., 1970. *Philosophy of Logic*. Upper Saddle River, NJ: Prentice–Hall.

Quine, W. V. O., 1976. Truth by convention. In *The Ways of Paradox and other Essays*, 2nd edn, pp. 77–106. Cambridge, MA: Harvard University Press.

Raatikainen, Panu, 2018. Gödel's incompleteness theorems. In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, fall 2018 edition. Available at: https://plato.stanford.edu/archives/fall2018/entries/goedel-incompleteness.

Raatikainen, Panu, Forthcoming. Neo-logicism and its logic. *History and Philosophy of Logic*.

Rawls, John, 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rayo, Agustin, 2007. Plurals. *Philosophy Compass*, 2: 411–427.

Rayo, Agustin and Williamson, Timothy, 2003. A completeness theorem for unrestricted first-order languages. In Jc Beall (ed.), *Liars and Heaps*, pp. 331–356. Oxford: Oxford University Press.

Rayo, Agustin and Yablo, Stephen, 2001. Nominalism through denominalization. *Noûs*, 35: 74–92.

Read, Stephen, 1979. Self-reference and validity. *Synthese*, 42: 265–274.

Read, Stephen, 2003. Logical consequence as truth preservation. *Logique and Analyse*, 183(4):479–493.

Read, Stephen, 2008. Harmony and modality. In C. Cédric Dégremont and H. Rückert L. Keiff (eds), *On Dialogues, Logics and Other Strange Things*, pp. 285–303. London: King's College Publications.

Read, Stephen, 2012. Relevant logic: A philosophical examination of inference. Available at: https://www.st-andrews.ac.uk/~slr/Relevant_Logic.pdf

Rescher, Nicholas, 1969. *Many-Valued Logic*. New York: McGraw-Hill.

Resnik, Michael, 1988. Second-order logic still wild. *Journal of Philosophy*, 85: 75–87.

Resnik, Michael, 1996. Ought there to be but one logic. In Jack Copeland (ed.), *Logic and Reality: Essays on the Legacy of Arthur Prior*, pp. 489–517. Oxford: Oxford University Press.

Resnik, Michael D., 1999. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press.

Resnik, Michael D., 2000. Against logical realism. *History and Philosophy of Logic*, 20: 181–194.

Restall, Greg, 1994. On logics without contraction. PhD thesis, University of Queensland.

Restall, Greg, 2002. Carnap's tolerance, meaning, and logical pluralism. *Journal of Philosophy*, 99: 426–443.

Restall, Greg, 2013. Assertion, denial and non-classical theories. In Edwin Mares Koji Tanaka, Francesco Berto and Francesco Paoli (eds), *Paraconsistency: Logic and Applications*, pp. 81–99. New York: Springer.

Ripley, David, 2012. Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, 5: 354–378.

Ripley, David, 2013. Paradoxes and failures of Cut. *Australasian Journal of Philosophy*, 91 (1): 139–164.

Ripley, David, 2015. Anything goes. *Topoi*, 34: 25–36.

Rossberg, Marcus, 2004. First-order logic, second-order logic, and completeness. In Vincent Hendricks, Fabian Neuhaus, Stig Andur Pedersen, Uwe Scheffler and Heinrich Wansing (eds), *First-Order Logic Revisited*, pp. 303–321. Berlin: Logos Verlag.

Rossberg, Marcus, 2006. Second-order logic: ontological and epistemological problems. PhD thesis, University of St Andrews.

Rossberg, Marcus, 2015. Somehow things do not relate: on the interpretation of polyadic second-order logic. *Journal of Philosophical Logic*, 44: 341–350.

Rossberg, Marcus and Cohnitz, Daniel, 2009. Logical consequence for nominalists. *Theoria*, 65: 147–168.

Rush, Penelope, 2014. Logical realism. In Penelope Rush (ed.), *The Metaphysics of Logic*, pp. 11–31. Cambridge: Cambridge University Press.

Russell, Gillian, 2016. Logical pluralism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2016 edition. Available at: https://plato.stanford.edu/archives/win2016/entries/logical-pluralism.

Sainsbury, Mark, 1991. *Logical Forms.* Oxford: Blackwell.

Scharp, Kevin, 2013. *Replacing Truth*. Oxford: Oxford University Press.

Schroeder, Mark, 2008. *Being For: Evaluating the Semantic Program of Expressivism.* Oxford: Oxford University Press.

Schwarz, Wolfgang, 2014. Against magnetism. *Australasian Journal of Philosophy*, 92: 17–36.

Sequoiah-Grayson, Sebastian, 2008. The scandal of deduction: Hintikka on the information yield of deductive inferences. *Journal of Philosophical Logic*, 37(1):67–94.

Shapiro, Lionel, 2011. Deflating logical consequence. *Philosophical Quarterly*, 61 (243): 320–342.

Shapiro, Lionel and Murzi, Julien, 2015. Validity and truth-preservation. In Lionel Shapiro and Julien Murzi (eds), *Unifying the Philosophy of Truth*, pp. 431–459. Dordrecht: Springer.

Shapiro, Stewart, 1991. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford: Oxford University Press.

Shapiro, Stewart, 1993. Modality and ontology. *Mind*, 102: 455–481.

Shapiro, Stewart, 2000. The status of logic. In Paul Boghossian and Christopher Peacocke (eds), *New Essays on the Apriori*, pp. 333–366. Oxford: Oxford University Press.

Shapiro, Stewart, 2006. *Vagueness in Context*. Oxford: Oxford University Press.

Shapiro, Stewart, 2014a. Pluralism, relativism, and objectivity. In Penelope Rush (ed.), *The Metaphysics of Logic*, pp. 49–71. Cambridge: Cambridge University Press.

Shapiro, Stewart, 2014b. *Varieties of Logic*. Oxford: Oxford University Press.

Sher, Gila, 1991. *The bounds of logic*. Cambridge, MA: MIT Press.

Sher, Gila, 1999. Is logic a theory of the obvious? *European Review of Philosophy*, 4:207–238.

Sher, Gila, 2008. Tarski's thesis. In D. E. Patterson (ed.), *Alfred Tarski: Philosophical Background, Development, and Influence*, pp. 300–339. Oxford: Oxford University Press.

Sider, Ted, 2010. *Logic for Philosophy*. Oxford: Oxford University Press.

Sider, Ted, 2011. *Writing the Book of the World*. Oxford: Oxford University Press.

Slater, B. H., 1995. Paraconsistent logics? *Journal of Philosophical Logic*, 24: 451–454.

Stalnaker, Robert, 1991. The problem of logical omniscience, I. *Synthese*, 89: 425–440.

Stalnaker, Robert, 1999. *Context and Content*. Oxford: Oxford University Press.

Stanovich, Keith E., 1999. *Who is Rational? Studies of Individual Differences in Reasoning*. Mahwah, NJ: Lawrence Erlbaum.

Steinberger, Florian, in press. Three ways in which logic might be normative. *Journal of Philosophy*.

Steinberger, Florian, 2017. The normative status of logic. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2017 edition. Available at: https://plato.stanford.edu/archives/spr2017/entries/logic-normative.

Stenning, Keith and van Lambalgen, Michiel, 2008. *Human Reasoning and Cognitive Science*. Cambridge, MA: MIT Press.

Stenning, Keith and van Lambalgen, Michiel, 2011. Reasoning, logic, and psychology. *Wiley Interdisciplinary Reviews. Cognitive Science*, 25:555–567.

Tarski, A., 1986. What are logical notions? *History and Philosophy of Logic*, 7: 143–154.

Tennant, Neil, 1994. Transmission of truth and transitivity of proof. In D. M. Gabbay (ed.), *What is a Logical System?*, pp. 161–177. Oxford: Oxford University Press.

Tennant, Neil, 2015. Logicism and neologicism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Available at: https://plato.stanford.edu/archives/win2017/entries/logicism.

Turner, Jason, 2015. What's so bad about second-order logic? In A. Torza (ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics and Language*, pp. 463–487. Cham: Springer.

van Atten, Mark, 2017. Luitzen Egbertus Jan Brouwer. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2017 edition. Available at: https://plato.stanford.edu/archives/win2017/entries/brouwer.

Varzi, Achille C., 2002. On logical relativity. *Philosophical Issues*, 12: 197–219.

Wansing, Heinrich, 2000. The idea of a proof-theoretic semantics. *Studia Logica*, 64: 3–20.

Wansing, Heinrich, 2006. Connectives stronger than tonk. *Journal of Philosophical Logic*, 35: 653–660.

Wansing, Heinrich and Shramko, Y., 2008. Suszko's thesis, inferential many-valuedness and the notion of a logical system. *Studia Logica*, 88: 405–429.

Wason, P. C., 1968. Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20: 273–281.

Weir, Alan, 1998. Naive set theory, paraconsistency and indeterminacy I. *Logique et Analyse*, 41: 219–266.

Weir, Alan, 2005. Naïve truth and sophisticated logic. In Jc Beall and Bradley Armour-Garb (eds), *Deflationism and Paradox*, pp. 218–249. Oxford: Oxford University Press.

Whittle, Bruno, 2004. Dialetheism, logical consequence and hierarchy. *Analysis*, 64 (4): 318–326.

Wilholt, Torsten, 2006. Think about the consequences! Nominalism and the argument from the philosophy of logic. *Dialectica*, 60: 115–133.

Williamson, Timothy, 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.

Williamson, Timothy, 2013. Logic, metalogic, and neutrality. *Erkenntnis*, 79 (Supplement 2): 211–231.

Williamson, Timothy, 2017. Semantic paradoxes and abductive methodology. In B. Armour-Garb (ed.), *Reflections on the Liar*, pp. 325–346. Oxford: Oxford University Press.

Wittgenstein, Ludwig, 1983. *Remarks on the Foundations of Mathematics*, 2nd edn. Cambridge, MA: MIT Press.

Wright, Crispin, 1986. Inventing logical necessity. In Jeremy Butterfield (ed.), *Language, Mind and Logic*, pp. 187–209. Cambridge: Cambridge University Press.

Wright, Crispin, 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.

Zardini, Elia, 2008. A model of tolerance. *Studia Logica*, 90: 337–368.

Zardini, Elia, 2011. Truth without contra(di)ction. *Review of Symbolic Logic*, 4 (4): 498–535.

Zimmermann, Thomas Ede and Sternefeld, Wolfgang, 2013. *Introduction to Semantics: An Essential Guide to the Composition of Meaning*. Berlin: De Gruyter.

Index