Trends in Logic 40

Andreas Kapsner

Logics and Falsifications

A New Perspective on Constructivist Semantics

Springer

Trends in Logic

Volume 40

For further volumes:
http://www.springer.com/series/6645

TRENDS IN LOGIC
Studia Logica Library

VOLUME 40

SCOPE OF THE SERIES

The book series *Trends in Logic* covers essentially the same areas as the journal Studia Logica, that is, contemporary formal logic and its applications and relations to other disciplines. The series aims at publishing monographs and thematically coherent volumes dealing with important developments in logic and presenting significant contributions to logical research.

The series is open to contributions devoted to topics ranging from algebraic logic, model theory, proof theory, philosophical logic, non-classical logic, and logic in computer science to mathematical linguistics and formal epistemology. However, this list is not exhaustive, moreover, the range of applications, comparisons and sources of inspiration is open and evolves over time.

Andreas Kapsner

Logics and Falsifications

A New Perspective on Constructivist Semantics

Springer

Andreas Kapsner
Munich Center for Mathematical Philosophy
Ludwig-Maximilians-Universität
Munich
Germany

ISSN 1572-6126 ISSN 2212-7313 (electronic)
ISBN 978-3-319-05205-2 ISBN 978-3-319-05206-9 (eBook)
DOI 10.1007/978-3-319-05206-9
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014935968

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgments

It is my pleasure to thank several persons and institutions who helped me in writing this book. It is a revised version of my doctoral thesis, which I wrote as a member of the LOGOS research group and defended in 2011 at the University of Barcelona, and which in 2012 received the Beth Dissertation Award from the Association of Logic, Language and Information (FoLLI).

The thesis was directed by Genoveva Martí, Sven Rosenkranz, and Heinrich Wansing. Their advice, questions, and encouragements have shaped this book. Also, the material benefited enormously from the input of my viva committee, which consisted of Hannes Leitgeb, José Martinez, and Graham Priest. Several others made lasting impacts. These include the members of the LOGOS group, scholars I met on research stays at the University of Melbourne, the Technical University of Dresden and the Munich Center for Mathematical Philosophy, and the audiences of various conferences in which the material was presented. I'd like to especially thank David Miller and Dave Ripley, who read the whole manuscript and gave a lot of stimulating feedback on it.

Funding of my research came from a FPI scholarship (BES-2006-11543) and Genoveva Marti's research projects HUM2005-00761 and FFI2008- 04263/FISO, financed by the Spanish Ministry of Science and Innovation, a fellowship from the Alexander von Humboldt foundation and the research project "New Logics for Verificationism" PI 1082/1-102288852134 funded by the German Research Foundation.

Some of the material has been published as Kapsner (2013) and is reprinted with permission. Early versions of some of the ideas appear in Pietz (2011 a, b).

Contents

1 Introduction 1
- 1.1 Analytical Table of Contents 2
- 1.2 Symbols and Abbreviations 7

Part I Background

2 Constructivism 11
- 2.1 Chapter Overview 11
- 2.2 Constructivism 11
- 2.3 Language, Logic, Metaphysics 13
- 2.4 Decidability 15
- 2.5 Undecidability, Bivalence, and the Law of Excluded Middle 17
- 2.6 Where to Start? 19
- 2.7 Truth Conditions. 20
- 2.8 Meaning as Use 20
- 2.9 Correct Assertibility or Truth? 22
- 2.10 Meaning Theory, Theory of Meaning and Semantic Theory 24
- 2.11 Revisionism versus Eclecticism 26
- 2.12 Metaphysical Conclusions 28
- 2.13 Chapter Summary 29

3 Intuitionism 31
- 3.1 Chapter Overview 31
- 3.2 Brouwer's Philosophy of Mathematics 31
- 3.3 Constructive Mathematics 32
- 3.4 The Axiomatization of Intuitionistic Logic. 33
- 3.5 Various Semantics 33
- 3.6 The BHK Interpretation. 34
  - 3.6.1 Correctness and Explosion. 38
- 3.7 Kripke Semantics 39
  - 3.7.1 Logical Constants 42

3.7.2 Logical Consequence 43
3.7.3 Kripke Semantics and the BHK Interpretation 44
3.8 Tensed or Untensed Provability? 47
3.8.1 Untensed Provability 48
3.8.2 Tensed Provability 49
3.8.3 Varieties of Tensed Provability 50
3.8.4 The Meaning of the Value 1 53
3.9 Chapter Summary 54

4 Gaps, Gluts and Paraconsistency 57
4.1 Chapter Overview 57
4.2 Relevant Logic 58
4.3 First Degree Entailment 59
4.3.1 Designated Values 61
4.3.2 Thinking Computers 62
4.4 Exactly True Logic 64
4.5 LP and K_3 66
4.6 Uses of Gaps 67
4.6.1 Presupposition Failure 67
4.7 Designated Values, Assertoric Content
and Ingredient Sense 69
4.8 Motivations for Paraconsistency 71
4.9 Dialetheism 72
4.10 Expressing the Law of Non-Contradiction 73
4.11 The Law of Non-Contradiction, Bivalence, and Duality 75
4.12 Analetheism 76
4.13 Chapter Summary 77

Part II Falsifications

5 From Proofs to Verifications and on to Falsifications 85
5.1 Chapter Overview 85
5.2 Stage I: Pure Verificationism 88
5.3 Stage II: Expanded Verificationism 91
5.3.1 Intuitionistic Logic? 92
5.4 Stage III: Hybrid Strategies 96
5.5 Stages IV and V: Expanded and Pure Falsificationism 99
5.6 Hybrid Strategies Again 101
5.7 Chapter Summary 102

6 Falsificationism 105
6.1 Chapter Overview 105
6.2 The Centrality of Incorrectness 106

6.3 Retracting, Surprising, and Misleading 108
6.4 What is a Falsification? 109
6.5 Correct Assertibility as Unfalsifiability 110
6.6 Example I: Legal Trials 112
6.7 Example II: Faultless Disagreement 113
6.7.1 General Strategies and Their Problems 113
6.7.2 Unfalsifiability and Analetheism 115
6.7.3 Agreeing to Differ 116
6.7.4 Falsificationism About Taste 117
6.8 Correctness and Truth 118
6.9 The Uncertain Truth Conditions of Conditionals 120
6.10 Falsificationistic Systems of Logic 121
6.11 Chapter Summary 121

Part III Logics

7 Stage Five: Pure Falsificationism and Dual Intuitionistic Logic 125
7.1 Chapter Overview 125
7.2 Dummett's Falsificationistic Logic 126
7.3 Falsificationistic Semantics 127
7.3.1 Rumfitt's "Safe Assertibility" 127
7.4 Kripke Semantics for Dual Intuitionistic Logic 128
7.4.1 Kripke Semantics: The Basic Setup 129
7.4.2 The Connectives 132
7.5 Some Characteristic Features of DIL 134
7.5.1 Paraconsistency 135
7.6 A BHK-Style Interpretation 137
7.6.1 Conjunction and Disjunction 137
7.6.2 The Conditional 137
7.6.3 Co-implication 142
7.6.4 Negation 143
7.7 Chapter Summary 144

8 Stage Two: Expanded Verificationism and the Logic N_3 145
8.1 Chapter Overview 145
8.2 BHK Interpretation 146
8.3 Kripke Semantics 147
8.3.1 The Connectives 149
8.4 Do We Want Gluts? 150
8.5 Features of the Logic N_3 153
8.5.1 Some Attempts to Get Contraposition Back 155
8.5.2 Embracing Contraposition Failure 157

8.6 Toggle Negation Versus Intuitionistic Negation 159
8.7 Chapter Summary 162

9 Stage Four: Expanded Falsificationism and the Logic N_{3f} 163
9.1 Chapter Overview 163
9.2 Falsificationistic Nelson Logics 164
9.3 Incoherence 167
9.3.1 Getting Rid of Incoherence 170
9.4 Safe Assertibility Reconsidered 172
9.5 Chapter Summary 173

10 Stage Three: Hybrid Strategies 175
10.1 Chapter Overview 175
10.2 Dummett on Hybrid Strategies 177
10.2.1 Gappy Semantics 177
10.2.2 Dummett Against In/Correctness Gaps 177
10.3 CV&IF: Additional Effects 180
10.4 Verificationism and Falsificationism Again (with Additional Effects) 181
10.5 A Tripartite Setup 182
10.6 Unifying the Account 184
10.7 Burden of Proof Distribution 185
10.7.1 Taking the BoP Outside of Legal Discourse. 186
10.7.2 The BoP in the Conversational Score 187
10.7.3 Summing Up the BoP Strategy. 189
10.8 Hybrid Consequence 190
10.9 Whatever Happened to Truth? 192
10.10 Chapter Summary 193

Part IV Summary

11 Summary 197
11.1 Resume 197
11.2 What is Constructivity? 199

Appendix A: Characteristics of Some of the Logics 203

Appendix B: Tableaux for N_3, N_{3f}, N_{AND}, $N_{AND\ f}$ and Hybrid Consequence 205

References 211

Index 215

Chapter 1
Introduction

The central themes of this book are certain non-classical logical systems, their philosophical motivation, and the meanings of their constants. My point of departure is the general constructivist line of argument that Michael Dummett has offered over the last decades. This argument expands to touch on a dazzlingly large number of important philosophical topics, but its root lies in Dummett's philosophy of language.

Even though most of his project is a modern version of verificationism, there is a clear strand of falsificationistic thinking in his writing that merits closer scrutiny. Indeed, Dummett himself has recently bemoaned the fact that this strand has not received much attention by his commentators.[1] This book explores where this largely untrodden path might lead.

Very succinctly put, the difference between the usual verificationistic picture and the alternative falsificationistic one is this: Understanding language consists in understanding under what circumstances an assertion would be correct. The verificationist story is that this consists in knowing under what circumstances an assertion would be *verified*. The falsificationistic story, in contrast, has it that one needs to know what would *falsify* the assertion, because an assertion will have to count as correct unless it is falsified.

One of the most interesting aspects of Dummett's new constructivism is that it seems to entail a revision of classical logic. For Dummett, the logic that should be adopted by a verificationist is intuitionistic logic, a logic that was developed in response to the philosophic views of L. Brouwer. Intuitionistic logic shows many characteristic differences compared to classical logic.

As will become clear in my discussion, I think that the central aspect of intuitionistic logic that makes it suitable for Dummett's verificationism is its rejection of the Law of Excluded Middle (henceforth LEM), while other peculiarities such as the rejection of Double Negation Elimination are quite accidental and dispensable features. The LEM says that, as a matter of logical validity, "*A* or not *A*" will always be assertible, no matter what sentence *A* stands for. For the verificationist, this is

[1] Auxier and Hahn (2007), p. 694.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_1,

only true if either "A" or "not A" can be verified, but this cannot be assumed to hold for every A, so the LEM must be rejected.

Now, one of the main claims I wish to make is this: If the LEM is what has to be given up in verificationism, then the move to falsificationism will have a different casualty, the principle of *Explosion*. This classically and intuitionistically valid principle, also known as *ex contradictione quodlibet*, tells us that we can infer whatever we wish from a contradiction: $A \wedge \neg A \vDash B$. Any logic that does not support such inferences is called a paraconsistent logic.

The normal way of supporting the principle of Explosion is this: A contradiction such as "A and not A" will never be true, no matter what sentence A we choose to plug in. An inference is valid iff (if and only if) the conclusion is true whenever the premises are true; therefore, the inference from something that can never be true to an arbitrary statement is always valid.

However, I claim that this kind of argument will not hold if we take the statement "A and not A" to be correctly assertible if it is not falsifiable, for the most natural way of giving falsification conditions for a conjunction will turn out to be the requirement that one of the two conjuncts is constructively falsified. But just as there was no guarantee that we can always verify an arbitrary statement or its negation, there is no guarantee that we can falsify either of them. Thus, given the falsificationistic account of what a correct assertion is, there is no guarantee that a contradiction will never come out assertible.

Falsificationism represents quite a radical departure from the usual verificationistic picture. I will spend quite some time in this book trying to get to grips with it. However, there are also more subtle ways in which falsifications can enter into a constructivist semantics. I will display the full spectrum of options and discuss the logical systems most suitable to each one of them. There are many forks on this path, and only one of them leads to intuitionistic logic.

The book is divided into three large parts. In the first part, important background information about Dummett's program, intuitionism, and logics with gaps and gluts is supplied. The second part is devoted to the introduction of falsifications into the constructive account. It turns out that there is more than one way in which one can do this. In the third part of the book, I detail the logical effects of these various moves. Below, I give a chapter by chapter overview of what is to come.

1.1 Analytical Table of Contents

Chapter 1 Introduction

This book examines the effects on logic that introducing the concept of *falsification* as a central notion of semantic theories will have. This introduction gives a first idea of what this might mean, and an overview of the following chapters.

Introduction to Part I: Background

The first part of the book consists of three chapters that will provide the foundation for the later discussion.

Chapter 2 Constructivism

This chapter, like the next two, is an introductory one that presents the critique of classical logic that Dummett put forward. It is grounded in his account of how meaning theories should be constructed, namely not in terms of mind-independent truth conditions, but rather in terms of proof or verification conditions. A semantic theory that gives the proof (verification) conditions of logically complex statements in terms of the proof (verification) conditions of their constituent statements is the basis of such an explication of meaning. The logical inferences that are licensed by this semantic theory are, arguably, those of intuitionistic logic.

Chapter 3 Intuitionism

The original ideas behind intuitionistic logic, its axiomatics, and two semantic theories are presented. Intuitionistic mathematics was the brainchild of L. Brouwer, who took mathematics to be about mental constructions, not abstract objects. These ideas were captured in a logical system by his student A. Heyting. I present the Brouwer–Heyting–Kolmogorov interpretation and the Kripke semantics for intuitionistic logic.

Chapter 4 Gaps, Gluts, and Paraconsistency

In this chapter, some semantical theories that allow for gaps and/or gluts are introduced, and the logics that can be based on them. In particular, I will present First Degree Entailment (FDE), strong Kleene (K3), and the Logic of Paradox (LP). An important idea that is introduced here is that logical consequence need not be defined as truth preservation, but might also be defined as non-falsity preservation. This makes no difference in some cases, in others it does. The notions of paraconsistency, dialetheism, and analetheism are introduced. Again, all this is important background material for the later chapters.

Introduction to Part II: Falsifications

This second part contains the most exegetical work. I try to analyze and systematize what Dummett has to say about the role of falsifications in semantic theories.

Chapter 5 From Proofs to Verifications, and on to Falsifications

I claim that the move from mathematical discourse to the empirical realm will have an influence on the logic that is motivated by the constructivistic semantic theory. This is because the intuitionistic explanation of negation is highly problematic in this setting. Dummett acknowledged that falsifications are necessary to fix the verification conditions of logically complex expressions. He then went even further and suggested that falsification might even be regarded as the central concept in a semantic theory and that logical consequence should transmit non-falsifiability.

To get some order into what Dummett offers us, I first give a clear account of the broad distinction I make between *verificationism* and *falsificationism*. The central tenet of verificationism is that an assertion is correct iff it is verifiable. The central idea of falsificationism, on the other hand, is that an assertion is correct iff it merely is not falsifiable.

Making more fine-grained distinctions, I then discern five stages of possible involvement of falsifications in a semantic theory. They are, in ascending order of falsificationistic predominance:

(I) pure verificationism, the view presented in the first two chapters leading to intuitionistic logic; (II) expanded verificationism, a verificationism (in the sense above) that uses falsifications to fix the meaning of complex statements; (III) hybrid strategies that rely both on verifications and falsifications in equal parts; (IV) expanded falsificationism, a falsificationism that relies on verifications to explain complex statements; and (V) pure falsificationism, the complete expulsion of all verifications.

These five stages, I will argue, all come with their own distinctive logics, and the last part of the book will go through these stages and their logics one by one.

Chapter 6 Falsificationism

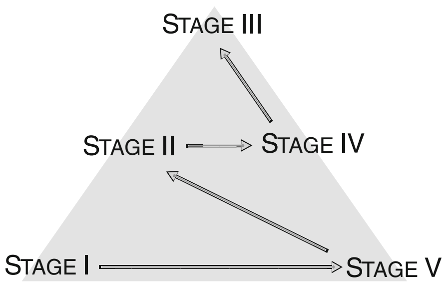
In this chapter, the idea of a falsificationistic theory of meaning is examined, without regard to the exact kind of falsificationism (pure or expanded) at issue. I further analyze Dummett's arguments for the idea that assertibility is nothing more than non-falsifiability, and I strive to give arguments of my own that make this position as plausible as possible. To this end, I give examples of areas of discourse that might be governed by a falsificationistic norm of assertion.

One of these examples concerns the assertions a defendant makes at a criminal trial. As most legal systems are working under the presumption of innocence, the defendant will be able to correctly make any assertion, as long as this assertion is not falsified. In other words, the prosecution has to bear the *burden of proof*, a concept that will feature prominently in the last chapter of Part III.

The second example I focus on is taste talk. If I say “Sushi is tasty”, then it is up to my audience to prove me wrong. If they cannot, my assertion will be correct, even if not all members of the audience are bound to like sushi. I will make a proposal what it would mean to prove me wrong in this case.

Introduction to Part III: Logics

In the last part of the book, I will present the logics corresponding to the different stages I mentioned above. As I will make plain, it will make sense to go through them in the following zig-zag pattern:



Chapter 7 Stage Five: Pure Falsificationism and Dual Intuitionistic Logic

As Stage I (intuitionism) has already been dealt with in Part I, I start this last part by presenting Dummett’s own proposal for a logic for Stage V: A logic that is only based on falsifications, the paraconsistent logic known as *dual intuitionistic logic*. This logic preserves not verifiability (as the concept of a verification is not utilized), but unfalsifiability. Choosing this property ensures that a speaker who asserts the premises will not incur further liabilities by asserting the conclusion. I present dual intuitionistic logic in a different semantical guise than Dummett did, which will make it more accessible. It will turn out, however, that this logic suffers from problems with complex statements, similar to those that intuitionistic logic had.

Chapter 8 Stage Two: Expanded Verificationism and the Logic N_3

In the first account that combines verifications and falsifications, I start with a preliminary discussion of how verifications and falsifications should be related. The upshot is that there are gaps (statements that are neither verified nor falsified), but no gluts (statements that are both verified and falsified). Based on this assessment, I will present the logic that most naturally arises. It is a species of the so-called Nelson logics. When it comes to conditionals and negations, there are some options to be explored, and I will go through the most important ones.

Chapter 9 Stage Four: Expanded Falsificationism and the Logic N_{3f}

Turning then to an account that pays tribute to Dummett's idea of falsificationism, I will show how to modify the logic of the previous chapter so that it transmits non-falsifiability. There are some worrying features of this new logic, as it seems to allow assertions that are quite incoherent. On the one hand, the logic is a paraconsistent one, which I claim is a good thing. This feature ensures the possibility for two persons to be correct, even though they contradict each other. However, it seems to allow the assertion of outright contradictions, and it does not satisfy modus ponens. I will then present a strategy to overcome these problems.

Chapter 10 Stage Three: Hybrid Strategies

I show in this chapter how verificationism and falsificationism can be combined. I give three main strategies (which might be used alongside each other): First, to simply differentiate areas of discourse in which assertions are correct iff verifiable, and others in which assertions are correct iff unfalsifiable. The second strategy is to open up a space for assertions that are neither correct (verifiable) nor incorrect (falsifiable). The last strategy is to make the norm of assertion dependent on the burden of proof. Paradigmatic here is legal discourse, where we might see the assertions of the defendant to be correct iff they are not falsifiable, and those of the prosecution as correct iff they are verifiable.

Chapter 11 Summary

I review the findings of the book and end by drawing some further philosophical conclusions.

1.2 Symbols and Abbreviations

A note on the logical symbols that will appear in this work: I will use $\neg$ to denote either classical negation or a generic unspecified negation (context will disambiguate). $\sim$ will stand for intuitionistic negation, $\multimap$ for dual intuitionistic negation, and $-$ for Nelson negation. I will only use two different symbols for conditionals, namely $\rightarrow$ for the material conditional and $\supset$ to denote a number of conditionals that are constructive in one sense or another (again, context will most of the time disambiguate sufficiently; if not, I will use subscripts as in $\supset_{\mathrm{TOL}}$).

I will employ the following abbreviations (for logical principles and often cited works of Dummett's):

DNE	Double Negation Elimination
DNI	Double Negation Introduction
EOI	Elements of Intuitionism (Dummett 2000)
LBM	The Logical Basis of Metaphysics (Dummett 1991)
LEM	Law of Excluded Middle
TOE	Truth and Other Enigmas (Dummett 1978)
TRUTH	Truth (Dummett 1978, pp. 1–25)
WTM	What is a Theory of Meaning (II) (Dummett 1993, pp. 34–94).

Part I
Background

Introduction to Part One

This first part of this book develops the necessary background for the developments in the second and third parts.

Chapter 2, entitled "Constructivism," introduces the basic philosophical outlook that underpins the following considerations. Mainly, but not exclusively, this book is inspired by Dummettian themes and ideas. I characterize his philosophy as a *constructive* one. This is not an unusual choice, but more often one will see him labeled an anti-realist. The reason why I choose to call him a constructivist instead is that I am more interested in his thoughts about logic and language than those about metaphysics, even though these areas are closely intertwined in his thinking.

Constructivism, as I shall define it, is a view on what suitable semantic values might be. A constructivist demands that these values need to be epistemically accessible, and this demand will have an effect on the admissible laws of logic. *Which* effect exactly that might be is the topic of the rest of the book.

Chapter 3, "Intuitionism," deals with the philosophical doctrine that goes under that name, and also with intuitionistic logic. In the semantics of intuitionistic logic we meet a first example of what a constructive semantics might look like. This will serve as the blueprint of the semantical theories that are to come, even if I will argue that the resulting logic need not be intuitionistic in each and every case. In fact, as soon as verifications and falsifications will take the place of proofs, I will present an argument to the effect that intuitionistic logic is a rather unlikely outcome. This argument, however, will unfold only quite a while later in the book.

Chapter 4, "Gaps and Gluts" introduces some themes that are not usually associated with Dummett's constructivism. However, it will become apparent in the course of the later parts that these ideas will present themselves quite naturally once we give up the idea that constructivist semantics has to follow strictly intuitionistic lines. In this preparatory chapter, however, constructive traits are not yet in the foreground. I discuss three sample logics that incorporate gaps (K3), gluts (LP), or both (FDE). In some form or another, all of these logics will make reappearances in the later parts.

As you would expect, most material in this first part is not overly original. However, I'm afraid that even for those acquainted with the topics, none of it can be skipped easily (though skimming and later backtracking might well do).

Dummett's philosophy is much like the proverbial elephant, and you will want to know whether this particular blind author is at the tusks, the knees, or even at the tail end of things. To find out, you will have to read Chap. 2.

Chapter 3 also introduces some of the semantical ideas underlying intuitionism in rather more detailed fashion than you would standardly find. Some choices that I make might even be somewhat controversial, though I should hope that I manage to motivate them well enough.

Chapter 4, "Gaps and Gluts" might be the one that is most dispensable for those who know about relevant, paraconsistent, and partial logics. Among the things that might even be new to such readers, I introduce *analetheism* as a lesser known alternative to dialetheism, and I present a novel take on the intuitive interpretation of first degree entailment.

Chapter 2
Constructivism

2.1 Chapter Overview

In this chapter, I will provide a brief summary of Dummett's constructivist program. My central topics, the meaning of the logical constants and the admissibility of logical laws, lie at the heart of a grand philosophical system, in which Dummett deftly strings together philosophical insights about language, logic and metaphysics. In some cases, it is quite impossible to understand his arguments about logical consequence without having at least a general idea of the outline of the whole program. This chapter aims to give such an outline.

The exegetical efforts needed to bring the central features of Dummett's difficult and extensive writing into a more or less streamlined form are considerable. I have no doubt that for many of the substantial claims I attribute to Dummett, some dissenting quote or other of his could be found. I believe that I give a fair and charitable version, but other equally charitable versions might look very different, especially if they lay more weight on other works than I do.[1]

A further point of this chapter, and in fact the one I shall address first, will be to fix some important terminology.

2.2 Constructivism

The first terminological item that deserves immediate comment is "constructivism" itself. Dummett, as we shall see, takes much inspiration from constructive mathematics and intuitionism. His contention is that the intuitionistic ideas, especially about the revision of the logical laws, can be translated to empirical discourse as well and help in giving a general account of meaning.

[1] For me, the central sources are his essays "TRUTH," "What Is A Theory Of Meaning (II)" (WTM) and the book "The Logical Basis of Metaphysics" (LBM).

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_2,

Intuitionism will be discussed in detail in the next chapter, but here is already a short summary of the position: There are no objectively existing mathematical entities, such as numbers or sets. Mathematics is therefore not about such abstract entities; rather, it is about mental *constructions*. Numbers and the like are products of the human mind, they are constructed by us.

It is not quite clear that Dummett's constructivism about empirical statements entails that the objects such statements talk about are constructed by us in a similar way. Calling his position "constructivism" is thus a slightly problematic (though standard) choice.

For reasons that we will soon come to see, a better choice would have been "verificationism." However, I need to reserve that term to contrast certain constructive theories that take verifications as the central concept in a semantic theory with others that take falsifications to be more important. Predictably enough, I will talk about "falsificationism" in the latter case. Verificationism and falsificationism will be specific examples of constructive theories, and it would be too confusing to use the first of these terms ambiguously. Bulky neologisms such as "veri/falsificationism" are ruled out for aesthetic reasons.

Maybe most commonly, Dummett's position is referred to as "anti-realism." However, to the uninitiated, it must be completely mystifying why a philosophical position on meaning and logic should bear such a name. Clearly, it is an apt name for a metaphysical doctrine.

Indeed, in a way the resolution of metaphysical questions might be seen as the ultimate goal of Dummett's thoughts about language and logic. It is his surprising and revolutionary idea that the only viable way to reach such a resolution leads through these seemingly unconnected topics.

The exact way in which language, logic and metaphysics are supposed to be related is quite complicated. I'll give an outline of these connections throughout the next sections. However, the metaphysical upshot of the constructive theories of meaning and logic is not my main concern in this book. Therefore, I will try to avoid the term "anti-realism" as a generic tag of Dummettian contentions. I will use it when metaphysical assumptions feature heavily in a line of argument, or when metaphysical consequences are more important than details about language or logic.

Here is a list of the terms I just mentioned, and what I will take these terms to mean. I only give you short slogans that aim to hit the core of the notions I have in mind. The exact meaning of these slogans will become clear in the course of this chapter.

CONSTRUCTIVISM The semantic values of statements must be, at least in principle, epistemically accessible. I will discuss two main kinds of constructivism: Verificationism and falsificationism.

VERIFICATIONISM A constructivism that takes the core of the semantic values to be a positive notion, namely verification conditions. Mathematical proofs are one particular species of verification, so that I will make it clear when I want to talk about empirical verifications only.

FALSIFICATIONISM A constructivism that takes the core of the semantic values to be a negative notion: *Falsification* conditions. Falsificationism in this sense is a relatively unknown species of constructivism, and it is one of the aims of this thesis to supply a fuller description. However, falsificationism will not make much of an appearance until Part II; the discussion in this first part is mostly about the better known verificationistic variations of constructivism.

INTUITIONISM By this, I will mostly mean the verificationistic theory of mathematical statements that I will present in more detail in the next chapter. Unlike the three items above, I will understand intuitionism to be tied to a very specific constructive logic, viz. intuitionistic logic. If the context makes it clear that mathematics is not specifically at issue, figures such as "the intuitionist" will simply be adherents of intuitionistic logic.

ANTI- REALISM Constructivism augmented by the claim that substantial metaphysical insights can be gained from the study of semantic values and logical laws.

So much for early terminological distinctions. Now let us immediately move on to the connections between constructivism and metaphysics alluded to in the last paragraphs.

2.3 Language, Logic, Metaphysics

Anti-realism, as is hard not to guess, is opposed to realism. Traditionally, under labels such as *idealism* and *platonism*, adherents of these stances had argued about what exists and what fails to exist. They fought over numbers, stones, mental states, quarks, and many things beside. Eventually, positivistic philosophers tired of these seemingly futile discussions and came to view them as quibblings over meaningless pseudo-questions. Their verdict was that philosophy had better turn its attention to the workings of language, in order to make sure that such meaningless propositions should immediately be debunked in the future.

One of Dummett's most spectacular ideas takes issue with this development: Yes, we should indeed concentrate on getting our philosophy of language right. Firstly, simply because language is important and interesting in its own right. But secondly, because linguistic insight will not *eradicate* metaphysical questions. Rather, such insight will make it possible to *answer* those questions.

However, it is not possible to answer those questions as they stand. The way they are phrased is, in Dummett's view, too *metaphorical*. About the specific example of mathematics, he writes:

> [W]e have here two metaphors: the platonist compares the mathematician with the astronomer, the geographer or the explorer, the intuitionist compares him with the sculptor or the imaginative writer; and neither comparison seems very apt (TOE, p. xxv).

Rather than asking whether a mathematician is like an astronomer or like a sculptor, that is, whether or not numbers exist on their own accord, Dummett asks what it is that makes our statements about numbers *true* or *false*. Or, if we are reluctant to talk about truth, what makes our statements *correct* or *incorrect*.

Realism, often relativized to a given area of discourse (mathematics, particle physics, ethics, etc.), holds that what makes the statements in this area of discourse true or false (numbers, quarks, moral facts) exists independently of anyone talking or thinking about it. Anti-realism, on the other hand, claims that the truth or falsity of those statements do depend on us and our abilities to *recognize* their truth values.

The first hard problem for those who are, unlike me, mainly occupied with the metaphysical questions is to decide whether this reformulation really captures the core of the traditional debates.[2]

Granting that Dummett indeed has hit on such a common core, the important next step is to focus on a particular question about the distribution of truth and falsity over the statements of the language. The question is this: Can we assume that every statement is either true or false? If we answer "Yes, we can," then we are assuming the principle of bivalence:

BIVALENCE Every statement is determinately either true or false.

Dummett claims that if we assume this principle, then we are realists about the area of discourse in question.[3]

In order to explain why bivalence should be doubted by an anti-realist, I will have to delve a bit deeper into the exact nature of the dependence of truth values on speakers and hearers that I mentioned above. The anti-realism that Dummett has in mind is not a relativistic doctrine of the "your truth is not my truth" kind, where anyone is free to attribute truth and falsity to statements in ways that are only constrained by their imagination.[4] The dependence is much more of a positivistic ilk: The anti-realist holds that a statement cannot be true unless we are, at least in principle, able to come to know of its truth. In mathematics, "to come to know of its truth" means to find a proof of a statement; in empirical contexts, it means to *verify* the statement.

This particular way in which truth values depend on us language users now has the following consequence: If we cannot assume that we can recognize the truth or falsity of every statement, then we cannot assume bivalence, either. And indeed this is an assumption that we should be reluctant to make if the talk we engage in is sufficiently sophisticated.

---

[2] See the introduction of Wright (1993) for detailed discussion on these and other matters that are only skirted here.

[3] "Realists about the area of discourse..." is a bit vague; the reason for this is that it is not always the existence of objects that is at stake, for example in discussions about the reality of the past or the future. In any case, even though Dummett often writes as if the case for or against realism has to be negotiated for each area of discourse separately, the arguments he brings forth are for the most extremely general. Therefore, I'll drop the relativization to a specific area of discourse until further notice.

[4] cf. Braver (2007) for a comparison between Dummett's and more liberal and anarchic conceptions of Anti-realism.

2.4 Decidability

To see why Dummett thinks bivalence cannot be assumed by a constructivist, we have to understand his notion of an undecidable statement. If it were not for undecidable statements, says Dummett, we would have no reason to doubt the validity of the principle of bivalence.

For the blueprint of the argument, we turn, once again, to mathematics. However, we immediately run into a problem: As we will see presently, Dummett's idea of decidability does not readily correspond to other conceptions of decidability found elsewhere in mathematics or logic, and what he tells us about his notion is far from perfectly clear. The best discussion I've seen of how to interpret Dummett in this respect is Shieh (1998), and this section draws heavily on that paper.

For Dummett, a mathematical statement is decidable if either we have already proved it (or its negation), or else if we know of a straightforward and sure-fire way of getting a proof of it or its negation. For the latter option, it is not enough to know of a method that will, in fact, decide the question; we need also to *know* that it will decide the issue, although we need not know whether the result will be positive or negative.

What makes this reading quite hard to come by is that it opens up the following possibility: Something that we are today able to prove might not have been decidable, even *in principle*, in the past. That is because in the past we might not have had the proof, nor any knowledge of how to construct it.

As was just noted, this conception has little to do with the standard idea of decidability. Consider our saying that propositional logic is decidable, whereas predicate logic is not. This has nothing to do with our *finding* decision methods. Propositional logic has always been decidable, and predicate logic will always remain undecidable in the classical sense. The question here hinges on the existence or non-existence of an algorithm, not on whether we know of that algorithm or not.

I will, however, be using the term in the way Dummett is using it. Here is an example to further illustrate this constructive sense of decidability. One fine Monday, an eminent mathematician wonders whether there are seventy[5] consecutive 7s in the decimal expansion of π. Although the mathematicians have computed π to a considerable length, they have not come across such a series yet. On the other hand, no one knows of a method of proving that such a series could not occur. The mathematician is thus dealing with an undecidable statement.

Now let us suppose that our mathematician, having no better idea how to tackle the problem, sets out to compute more and more decimal places of π. On Thursday, she comes across seventy consecutive 7s. She has obviously proven the statement she was wondering about, unless she made a mistake in the calculation. Therefore, it is equally obvious that the statement is now decidable. But this does *not* change anything about the fact that the statement was *undecidable* on Monday. The mathematician has not found out that the statement was decidable all along.

---

[5] The example is adapted from Brouwer, only that he used *seven* 7s. These have in the meantime been found in the decimal expansion of π. As far as I know, seventy 7s have not yet turned up.

This idea of undecidability appears to be a very strict one, then. This seems especially true in view of the idealizations the constructivists are ready to make by talking about provability or constructibility *in principle*. Such talk is necessary because the constructivist would not like to be confined to what he has *actually* constructed. For example, if no one had ever thought about the number 347867536893243, we still would like to count it as one of the entities that are available to us, because we have a straightforward way of constructing it out of the basic building blocks of mathematics. Likewise, the question whether this number is prime or not is decidable, because we know of algorithms that we could employ here.

Even if the number should be so large that it would take us impossibly long to write it down, we would like to gloss over that fact and push that impossibility aside with the remark that we are only dealing with constructibility "in principle." Even more strikingly, the intuitionists have no problem with extending the range of "in principle constructible" entities to the transfinite.

But then, one might ask, why is the case different with the construction of the sequence of the decimal expansion of π? Surely, we would like to say that the question was "in principle" decidable even on Monday! The idealization involved here is, after all, a much lighter one than the one involved in getting us to a number so large we could never hope to write it down in all our lifetimes: all the mathematician had to do was calculate for three more days.

This reasoning is mistaken because our mathematician had no *guarantee* that she would come across seventy 7s, no matter how long she would have searched. The problem is not that the seventy 7s might have been so far down the line that it would have taken an impossibly long time to compute that far. The crucial point is rather that she could not have rested assured that she would hit the seventy 7s, no matter how much time she was willing and able to spend on the search.

A good way of putting the matter is this: The notion of a decidable mathematical statement is, in Dummett's use, a completely *epistemic* concept. Its extension solely depends on the state of the art in mathematics.

In addition to that, we will see in the next chapter that those constructivists who adopt intuitionistic logic will not be able to say that there are *absolutely* undecidable statements, i.e., statements which we will never be able to prove and never be able to disprove.[6] Undecidability is thus for the intuitionist always *pro tempore*, that is, the most an intuitionist can mean when she says that a statement is undecidable is that *at the present moment* there is no known decision procedure.

[6] There is surely evidence in Dummett's writings that he sometimes understands "*in principle* undecidable" in this sense. But on balance, I think he more often uses the phrase "in principle" in the sense outlined above. See for example WTM p. 45, where he talks about "sentences which are, in practice or even in principle, decidable, that is, for which a speaker has some effective procedure which will, in a finite time, put him into a position in which he can recognize whether or not the condition for the truth of the sentence is satisfied."

2.5 Undecidability, Bivalence, and the Law of Excluded Middle

In dealing with an undecidable statement, we have nothing that guarantees that we can come to know of its being true or false. But lacking such a guarantee, a constructivist does not feel entitled to claim that the statement is either true or false. It can only be so if there is a way to find out whether it is true or false, and we have no guarantee that there is such a way.

This talk of guarantees sounds a bit stiff; a constructivist might rather want to say that there may be statements for which we can find no proof or disproof.[7] But, as I just mentioned above, the intuitionist cannot refer to such absolutely undecidable statements, even if he wanted to. This problem (if it is a genuine problem) is caused by the way in which intuitionists explain negations, as we'll see in Sect. 3.6.

This might well be a problem for Dummett also. As I mentioned before, Dummett thinks that anti-realists should turn away from classical logic and toward intuitionistic logic.

To give up bivalence means already to give up an important part of classical logic. The classical truth table semantics of the logical connectives presupposes bivalence. There *might* be an alternative story that does not invoke bivalence and still motivates all inferences of classical logic.[8] The much more likely outcome, however, is that without bivalence, some inferences will lose their validity.

This is exactly what happens in intuitionistic logic, which will come under closer scrutiny in the next chapter. However, let me already point to some characteristic features of this logic.

The first and most famous feature is the failure[9] of the Law of Excluded Middle (LEM):

$$\text{LEM} \ \vdash A \vee \neg A$$

Even those unacquainted with intuitionistic doctrine will have no trouble seeing why this law must be dubious if bivalence cannot be assumed. If there is a problem at all, it might be to appreciate that there is a *difference* between bivalence and the LEM at all. But there is: Bivalence is a doctrine about semantic values, while the LEM records that, as a matter of logic, "A or not A" will always hold.

In general, it is important to keep semantic principles and logical laws apart; I will try to aid the reader by writing the names of semantic principles such as "bivalence" in lower case letters and letting the names of logical laws start with upper case letters. Also, I will use abbreviations such as "LEM" only for logical laws.

---

[7] I use "disproof" to mean the proof of the negation of a statement.

[8] Ian Rumfitt has suggested such a strategy in Rumfitt (2007).

[9] Like most propositional non-classical logics, intuitionistic logic is strictly weaker than classical logic in that classical logic validates all intuitionistic inferences. Non-classical logics are often described in an impressionistic way by pointing out which kind of classical inferences are not supported, as in this section.

Marking the difference between semantic principle and logical law does not change the fact that the connection between bivalence and LEM seems perfectly straightforward.[10] Such lucid plausibility can hardly be claimed for the second most famous characteristic feature of intuitionistic logic, the failure of Double Negation Elimination (DNE):

$$\text{DNE} \quad \neg\neg A \vdash A$$

To understand this failure, one needs to understand the way in which negation is explained in intuitionism. Without this explanation (for which you will have to wait until the next chapter), the impossibility to infer "A" from "not not A" must seem quite bizarre. To make matters even more puzzling, consider that Double Negation Introduction (DNI)

$$\text{DNI} \quad A \vdash \neg\neg A$$

is actually *valid* in intuitionistic logic, as is the following (nameless) inference:

$$* \quad \neg\neg\neg A \vdash \neg A$$

clearly, there is something very strange going on with double negations in intuitionistic logic. This can also be seen by considering the following inference, called the Law of Excluded Third (LET). Unlike LEM, this law is, again, *valid* in Intuitionistic logic.

$$\text{LET} \quad \vdash \neg\neg(A \vee \neg A)$$

Just as LEM corresponds to bivalence (i.e., "every statement is either true or false" [11]), this logical law corresponds to the semantic principle that Dummett calls *tertium non datur*: "No statement is neither true nor false." [12]

And just as Dummett wants to deny LEM and accept LET, he wants to deny bivalence and accept tertium non datur. That is, he is committing himself to the claim that no statement is neither true nor false, but not to the claim that every statement is either true or false. As he acknowledged at one point, "this confused some readers."[13] I dare say it continues to do so.

As I said, all this will become somewhat more perspicuous once the intuitionistic explanation of negation is in place. However, even then the mysteries will not disperse

---

[10] To spell it out, the connection is established by an appeal to the disquotational scheme

"A" is true iff A,

the principle that a disjunction is true iff at least one of the disjuncts is and the principle that $\neg A$ is true iff A is false, and the semantic demand that no statement may have more than one truth value. None of these principles is completely uncontroversial, but they surely are intuitively plausible.

[11] TOE, p.XX.

[12] ibid.

[13] ibid.

completely. I will eventually, in the last part of this book, come to suggest constructive logics in which all double negation laws hold.[14]

2.6 Where to Start?

Instead of getting any further into the details of the logical systems that lie ahead, let us zoom out again to get the big picture of Dummett's philosophy back into view. Until now, we have seen the connection he forged between positions in metaphysics (realism vs. anti-realism), philosophy of language (bivalence or not) and logic (classical vs. intuitionistic). If this connection is really as stable as he claims, then conclusive arguments for one of these positions may lead to solutions in the other areas. But where exactly is it most likely that such conclusive arguments may be found?

As I said above, Dummett agrees with his positivistic predecessors that the metaphysical dispute as traditionally conducted shows little promise of providing a secure first foothold. This debate only presents us "with alternative pictures. The need to choose between these pictures seems very compelling; but the non-pictorial content of the pictures is unclear".[15]

Maybe we should start with logical considerations, then. However, as purely formal and uninterpreted systems, there seems little wrong with either classical or intuitionistic logic. It is the interpretation of the logical vocabulary that, if anything, will make a significant dent in this discussion. In other words, the question is what the logical constants might mean. [16]

The natural place to start, then, is the analysis of language; indeed, Dummett holds that the analysis of language supplies the base to all of philosophy:

> [T]he theory of meaning is the fundamental part of philosophy which underlies all the others. Because philosophy has, as its first if not its only task, the analysis of meanings, and because, the deeper such analysis goes, the more it is dependent upon a correct general account of meaning, a model for what the understanding of an expression consists in, the theory of meaning, which is the search for such a model, is the foundation for all philosophy (...) (Dummett 1973, p. 669).

This quote comes from Dummett's famous book on Frege's philosophy of language. He has appropriated many of Frege's insights, such as the idea that words only have meaning as parts of whole statements and that the meaning of utterances in other moods is derivative of utterances in the assertoric mood.

---

[14] The status of LEM and LET will then depend on whether we are dealing with a verificationistic or a falsificationistic theory. In the verificationistic case, both will fail, and in the falsificationistic case, both will hold.

[15] LBM, p.10

[16] There *are* very important proof theoretic arguments for intuitionistic logic. These arguments, however, are also seen as concerning the *meanings* of the logical constants. Indeed, the field of study these arguments fall into is now known as *proof theoretic semantics*. Important as this field is, I will not consider it further.

The central questions for Dummett then are these: What do our assertions mean? What is it to understand a statement, what is it for it to be true? If these really are the fundamental questions of *all* of philosophy, then they are certainly also the questions we should start with in the slightly more circumscribed investigation of language, logic, and metaphysics.

2.7 Truth Conditions

A widely accepted way of specifying the meaning of statements is to equate them with their *truth conditions*. To understand a statement is to apprehend under which circumstances it would be true.

One may think that this strategy would seem unpalatable to the Dummettian constructivist, especially in view of Dummett's earlier work. There, he advised to steer clear of truth conditions in the theory of meaning and to employ *assertibility* conditions instead. I will come back to this anon; let us for now concentrate on the later[17] position, in which Dummett offers that all parties can agree on the use of truth conditions, but that their differences will become apparent when they spell out the properties of the truth conditions alluded to.

For the realist, these truth conditions will be independent of any means to come to know them. Their notion of truth is *epistemically unconstrained*.

The Dummettian constructivist will object to such notions of truth and truth conditions. However, the objection will not turn on the pictorial and unclear content of his metaphysical assumptions. Rather, the constructivist claims that the use of epistemically unconstrained truth conditions gives an unsatisfactory account of meaning.

2.8 Meaning as Use

Here is a sketch of Dummett's line of argument against epistemically unconstrained truth conditions.

How, he asks, can such a truth condition be constitutive of the meaning of a statement that people are supposed to be able to use competently? If the obtaining of a truth condition is systematically beyond my grasp, how could I possibly come to understand it?

Take a non-mathematical example, adapted with slight modification from Dummett: Suppose that a certain chap, Jones, has up to now been leading a sedate life

[17] The talk of "earlier" and "later" is useful, but somewhat problematic. The change from assertibility to truth conditions in Dummett's thinking is not an abrupt one; rather, one sees a steady increase in the statistical likelihood of his talking about truth conditions, with the tipping point probably somewhere in the late 1970s. More on that in Sect. 2.9.

without encountering any particular perils. Now consider the statement "Jones is brave." If we assume bivalence, then this statement is true or false.

But we cannot *tell* whether Jones is brave or not until we watch him getting himself into a dangerous situation. Only then will his bravery or cowardice become apparent. Of course, there *might* be brain structures that predispose him already now to respond bravely in face of danger. Even if this were so, as of now we do not know how to find out about bravery by checking brain scans. If we assume that we are not in a position of putting Jones to the test by menacing him somehow, then we have no way to find out about the truth or falsity of the statement "Jones is brave."

Still, if we assume bivalence, the described circumstance is either among the conditions for the truth of that statement, or else among the conditions under which "Jones is not brave" is true.

If we accept this, Dummett argues, we cannot fully grasp the meaning of our statements. The meaning of the statement is made up of truth conditions, the present circumstance is one in which either "Jones is brave" or "Jones is not brave" is true, but we cannot know which. If we were grasping the meanings of the two statements fully, then we would have to be able to tell which statement is true under the present conditions.

But if we cannot be sure that we can grasp all aspects of the meaning of a statement, then the question becomes this: How are those aspects that are beyond our grasp helping to explain how linguistic exchanges function?

According to Dummett, it makes no sense to have a concept of meaning that has features that are not needed to describe linguistic behavior. This behavior is regulated by our notions of *correct assertibility*. If two speakers agree about all the ways in which a statement can correctly (and incorrectly) be used, then they attach the same meaning to the statement.[18] An aspect of a truth condition that is firmly out of our epistemic range cannot make a contribution to an agreement on whether a given statement was used correctly or not, and has consequently no place in a theory of meaning.

The above is Dummett's take on the oft-repeated Wittgensteinian insight that "meaning is use": If there are aspects of meaning that we can never be sure speakers grasp, no matter their linguistic behavior, then our meanings are too finely grained. In that case, our "theory of meaning is left unconnected with the practical ability of which it was supposed to be a theoretic representation."[19] A full grasp of meaning must be able to be demonstrated or *manifested* in one's use.

Epistemically unconstrained truth conditions fail this requirement of manifestability. The matter, says Dummett, stands differently if truth conditions are epistemically *constrained*. If it is always in principle possible to come to know them, then a manifestation of understanding will always be possible.

Take Jones again. What shows that you understand what "Jones is brave" means is that you recognize that certain behavior in a dangerous situation is brave. That is,

---

[18] And this is the reason for his earlier view that assertibility conditions suffice for specifying the meaning of a statement.

[19] WTM, p.71

you know what would verify the statement: A certain type of behavior in a dangerous situation.

If we want to explain meaning in terms of truth conditions, then the concept of truth should not contain more than what is needed: A statement is true iff it can be verified, and to know the meaning of a statement is to know the conditions under which we would consider it verified. Such a theory of meaning, Dummett claims, is in a much better position to explain the role meaning plays not only in language *use* but also in language *acquisition*.

That the latter is a problem for the realist can be seen if we ask how we can learn to speak if the correct meaning of the statements we are presented with is beyond our cognitive reach. We learn to speak and to understand language by being trained in circumstances that we are able to recognize, not in circumstances that go beyond our epistemic grasp. It is then quite unconceivable how we come to know any parts of meaning that treat on unrecognizable features of reality.

These then are the famous twin challenges that Dummett brings forth against realistic semantic theories: The manifestation challenge and the acquisition challenge. He has expressed these challenges many times, often with subtle differences in the presentation. [20]

Furthermore, there is an extensive literature on the question how successful these challenges really are.[21] I will not even attempt an overview of the ramifications of this discussion. Instead, I will now come back to the above-mentioned switch from assertibility conditions to truth conditions.

2.9 Correct Assertibility or Truth?

In this section, I'll try to track, rather roughly, Dummett's meanderings back and forth between two core tenets:

1. The content of an assertion is given by those circumstances in which an assertion of it would be correct. The concept of truth is not really needed in a theory of meaning.
2. The content of an assertion is given by its truth conditions. The notion of truth is epistemically constrained. To make a correct assertion is simply to assert something true.

For Dummett, the general trend over the years has been to move away from position 1 and toward position 2. However, he often went back and forth between these positions, sometimes marking the difference, sometimes not. To convey a sense of this development, here are some salient quotes:

First up, a quote from his famous early paper, "TRUTH" (1959):

---

[20] Apart from the sources I cited, Dummett (1975) offers a crisp presentation of the two arguments.

[21] One might start an exploration here: Hale (1999), Miller (2002), Miller (2003), Rosenkranz (2002), Tennant (2002) and Wright (1993).

> [W]e should abandon the notion of truth and falsity altogether. (...) We no longer explain the sense of a statement by stipulating its truth-value in terms of the truth-values of its constituents, but by stipulating when it may be asserted in terms of the conditions under which its constituents may be asserted (TRUTH, p.17).

The reprint of this essay is preceded by an extensive comment in the preface of TOE (1978). This comment seems to mark a decisive move from position 1 to position 2:

> On the way of putting it I adopted, one first proposes explaining meaning, not in terms of truth, but in terms of the condition for correct assertion, and then declares that, for statements whose meaning is so explained, the only admissible notion of truth will be one under which a statement is true when and only when we are able to arrive at a position in which we may correctly assert it. But, in that case, it would have been better first to state the restriction on the application of 'true', and then to have held that the meaning of a statement is given by the condition for it to be true in this, restricted, sense of 'true'. (...) Thus I should now be inclined to say that, under any theory of meaning whatever (...) we can represent the meaning (...) of a sentence as given by the condition for it to be true, on some appropriate way of construing 'true': the problem is not whether meaning is to be explained in terms of truth-conditions, but of what notion of truth is admissible (TOE, p.xxii).

Although this passage suggests a clear break with the old position, one will often find passages in later writings of Dummett that sound more like the original proposal, such as this passage from LBM (1993):

> It is plain that any account of the practice of assertion will supply us with a general notion of the correctness and incorrectness of assertions. The root notion of truth is then that a sentence is true just in case, if uttered assertorically, it would have served to make a correct assertion. (...)
>
> Thus the content of an assertion is taken as determined by the condition for it to be correct, and this in turn is identified with the condition for the sentence to be true: [W]e know what has been asserted when we know in what case the assertion is correct (LBM, p. 165–166).

In sum, it seems fair to say that Dummett thinks that there is a very close connection between truth and correct assertibility. However, the above quotes show that he is rather shaky on the question which of them has priority in an account of meaning. [22]

In addition, we have seen that he usually sees those two notions closely tied to the notion of verifiability: An assertion of a statement is correct and the asserted statement is true iff it is verifiable.

Nonetheless, there are also considerations of his under which the triad of truth, assertibility, and verifiability becomes undone in various ways, some of which I am going to discuss in quite some detail in the following chapters. In Sect. 3.8, I will discuss more thoroughly what the close tie between verifiability and truth brings us. Much later, in Part II, we will see what happens if correct assertibility is tied, not to verifiability, but to *non-falsifiability*. In that case, it will turn out that the close tie between correct assertibility and truth might not be plausibly upheld.

Whenever correctness and truth come apart, I will in principle stick with the earlier, i.e., the first of the two tenets at the beginning of this section: Essential to a grasp

---

[22] Cf. Kirkham (1995), p. 248 ff. and Green (2001), p. 24. for further discussion and more sample quotes.

of meaning is a grasp of correct assertibility conditions. From that basis, however, I will keep an interested eye on what might be said about truth, e.g., whether it might still be argued to correspond to correct assertibility in some way, whether it should be altogether abandoned by a constructivist, etc. However, until the falsificationistic theories of Part II come along, the identification of truth and correct assertibility is less controversial (though still far from unproblematic).

2.10 Meaning Theory, Theory of Meaning and Semantic Theory

Let us take stock: Up to now we have a rough idea of how the constructivist conceives of meaning. I have also, though not yet in any detail, told you that under this conception some classical laws of logic might become problematic.

This section aims to spell out the transition from a given theory of meaning, whether constructive or not, to a notion of logical consequence more clearly. This will be the general blueprint of what will happen in many of the following chapters, starting with the very next one, in which the case for intuitionistic logic is examined.

To get started, we need once again to clear up some Dummettian terminology that could otherwise generate confusion:

- A MEANING THEORY for a particular language records all that a speaker needs to know, whether explicitly or implicitly, in order to be considered a competent speaker of that language. In contrast,
- A THEORY OF MEANING gives the general form in which a meaning theory has to be presented. "The task of a theory of meaning is to give an account of how language functions, in other words, to explain what, in general, is effected by the utterance of a sentence in the presence of hearers who know the language to which it belongs."[23]
- A SEMANTIC THEORY. This is a theory of how the correct assertibility (or the truth) of a statement is determined by its semantic value, and how its semantic value depends on the semantic values of its parts.

The idea of semantic values and their composition basically goes back to Frege. The semantic theory will have to answer questions about the reference of singular terms and the like. As we are mainly interested in questions of logic (and are going to concentrate on the propositional case), the question that we need to concentrate on is this: How do logically complex statements receive their semantic values? Assuming compositionality, how does the semantic value of a logically complex statement depend on the values of the constituent statements? For example, if we know the

[23] LBM, p.21. Maybe "theory of language use" would have been a better label. Not only easier to discern from "meaning theory," it would have made the rest of this quote sound less puzzling: "The notion of meaning itself need not, therefore, play an important role in a theory of meaning." (ibid.) It might not, if there were no tight connection between meaning and use. However, we have seen in Sect. 2.8 that Dummett thinks that such a tight connection does hold.

value of A and the value of B, how do we get to the value of "A and B," "If A, then B" and so on?

This is the point at which it will be decided whether our theory is a constructive one or not. If we choose semantic values that are epistemically accessible (e.g., proof conditions, verification conditions, falsification conditions), then we are setting up a constructive semantics.

Once we have found out the semantic value of a statement, the next task of the semantic theory is to tell us whether the statement is correctly assertible/true in a given state of affairs or not. With this information, we can finally move on to a conception of logical consequence: An inference will be valid if it never fails to transmit correct assertibility/truth from premises to conclusions.

Let us see by way of example how the realistic semantic theory that gives the meaning of the classical logical constants works. It is as well known as it is simple. There is no mystery in how the semantic value of a statement relates to truth or falsity: The semantic value of a statement simply *is* its truth value.

The determination of the truth value of a complex statement is completely taken care of by the truth tables of classical propositional logic and the assumption that every statement takes exactly one of the truth values "True" and "False." The truth value of negations and disjunctions can be computed with the aid of the following tables:

$\neg$	
T	F
F	T

$\vee$	T	F
T	T	T
F	T	F

Given the usual definitions of the other connectives and, crucially, the assumption of bivalence, this is enough to give us full classical logic.

We have seen that in a constructive semantic theory, on the other hand, no assumption of bivalence can be made. Consequently, the constructivist claims, the account of the semantic values of complex statements and of logical consequence will have to be a different one.

Dummett's prime example of a semantic theory that is more adequate than the classical one is, once again, the account of mathematical statements in Brouwer's intuitionism: Such a statement is correctly assertible/true only if it is provable, and we cannot assume the correct assertibility/truth of a theorem if we cannot give a proof, or at least a method of constructing such a proof. The meanings of the logically complex statements are in turn given in terms of proofs of the constituent statements. In the next chapter, I will introduce two semantic theories that have been argued to be suitable for constructive theories. These will be the Brouwer-Heyting-Kolmogorov (BHK) interpretation[24] and the more formal Kripke semantics[25] for intuitionistic logic. The first spells out what a proof of a complex statement is by telling us, for example, that a proof of a conjunction is composed of two proofs, one for each

[24] see Sect. 3.6.

[25] see Sect. 3.7.

conjunct. The latter is a version of Kripke's well-known possible worlds semantics for modal logic.

2.11 Revisionism versus Eclecticism

But again, before getting into the details, let me try to get the big picture back into view. A meaning theory will contain all that a speaker must know in order to qualify as a competent speaker of a specific language. A theory of meaning outlines the general features of such specific meaning theories and is informed by the underlying semantic theory. The rules and laws of logic, finally, are validated by that semantic theory.

Given all this, we can now address the following question: How can a logical practice be criticized and revised by a theory of meaning that purports to be based on the idea that "meaning is use"? [26] If meaning is use, then isn't any use of statements that people make in a systematic way bound to be correct? In particular, where this "use" is the drawing of inferences, isn't each and every inference a community accepts going to be constitutive of the meaning of the logical vocabulary and therefore immune from revision? Isn't it enough to point to the fact that we are happy to infer A from $\neg\neg A$ to establish the validity of that inference?

Here is Dummett's answer, couched in the terms we have introduced over the last sections: Suppose that the semantic theory that backs up a logic arises from a theory of meaning that can be criticized on systematic grounds, such as the arguments from manifestation and acquisition I sketched above. In that case, that criticism is transmitted, via the semantic theory, to the logic itself. Then, even if it might be argued that the logic is well entrenched, in that the speakers themselves will not hesitate to draw inferences in accordance with its rules, this practice can be criticized by the theoretician. The practice on the whole, they will contend, is incoherent, and should be revised so as not to lead us astray in our reasoning.

The next task will then be to supply a package of semantic theory and logic that is better equipped to answer the worries that have been raised about the theory of meaning. Presumably (I am not following anything Dummett explicitly says here), if it should occur that there are several candidates, some principle of *minimum mutilation* should lead our choosing. A logic that is based on a sound linguistic basis and preserves more of the inferences people are apt to draw should then be preferred to one that suggests more bizarre revisions. (I stress this because I believe I can later present an alternative to intuitionistic logic that, among other virtues, is closer to the inferences we would normally like to draw.)

But is revision inevitable as soon as we find that a constructive logic can be underwritten in the way sketched above?

According to Dummett,

> [i]ntuitionism (...) raises two philosophical questions.

[26] Dummett is fully aware that at this point he can't claim to be following Wittgenstein any more, cf. LBM p. xi.

> 1. Do intuitionists succeed in conferring a coherent meaning on the expressions used in intuitionistic mathematics, and, in particular, on the logical constants?
>
> 2. Is there a ground for thinking that classical mathematicians *fail* to confer an intelligible meaning on logical constants, and on mathematical expressions in general, as they use them? (EoI, p. 251).

Only a positive answer to *both* questions will lead to a revision. Nonetheless, Dummett, somewhat unenthusiastically, acknowledges that one *might* answer the first question positively and the second negatively. That is, one might say that the meaning of the connectives are, in the same area of discourse, equally intelligible under the intuitionistic and the classical interpretation.

Granted, such "eclecticism"[27] is of no great interest to anyone who hopes to find answers to metaphysical questions. The only route from the meaning of logical constants to anti-realistic metaphysical conclusions is to hold that constructive logicians give coherent meaning to its constants, while classical logicians *do not*.

But logic and its philosophy are of interest in itself, even if its study does not entail metaphysical conclusions. Let us try to get an idea of what it might be like to accept the intelligibility of both intuitionistic and classical logic.

The most obvious way in which such an eclectic position could be construed is as a pluralism of languages, much in the spirit of Carnap's principle of tolerance.[28] The logical constants of classical logic and those of intuitionistic logic are two distinct sets of constants. One natural thing to say is that one employs the intuitionistic language when one is concerned with the transmission of *provability* or *verifiability*, while one turns to the classical language when epistemically unconstrained *truth* is at issue.

If this policy is adopted, then an important question will be on what side assertibility will come down: Will an assertion be correct iff it is true, will it be correct iff it is verifiable or something else altogether? Here, we obviously see one of the cases in which Dummett's firm connection between truth, verifiability, and assertibility could dissolve.

A new kind of logical eclecticism that offers a way of viewing *both* intuitionistic and classical logicians as concerned with truth and using the same language has been suggested by J.C. Beall and G. Restall.[29] They call their new position *logical pluralism*.

Their thesis is that both intuitionistic and classical inferences are valid iff they preserve truth in every possible case. The notion of a "case" is intentionally underspecified. It is in precisifying this notion that the different logics come about. If a case is understood to be a complete and consistent *world*, we get classical logic.

If, on the other hand, we consider cases to be *constructions*, the resulting logic will turn out to be intuitionistic logic. A construction is very much what I will call

[27] EOI, p. 250

[28] Carnap (1959)

[29] Beall and Restall (2006)

a "stage of investigation" when I discuss the Kripke semantics of intuitionistic logic in Sect. 3.7.[30]

According to Beall and Restall, what kind of "case" we take the present circumstance to be is a question of our interest. The different cases do not generally represent different areas of discourse.

2.12 Metaphysical Conclusions

Whether the eclecticist chooses the traditional route of Carnapian tolerance or the modernized version of Beall and Restall, the following seems hard to evade: In logic, there will be no morals, and very little metaphysics. Presumably, the eclecticist who endorses both classical and intuitionistic logic as intelligible will simply be committed to the same metaphysical upshots[31] as the monogamous classicist.

I am more concerned with the first of the questions Dummett raised at the beginning of the last section than the second one. That is, I am more interested in the question whether intuitionistic logic is really backed by a constructivist theory of meaning than in the question whether a realistic theory of meaning is coherent. As we have just seen that this might raise worries about the metaphysical fruits such an enterprise might bear, let me end this chapter by making my policy in this respect clear.

The connection between language, logic, and metaphysics is fascinating and puzzling in equal measures. I believe that much of the enduring appeal that has led many to study Dummett's very difficult writings lies in his bold reconception of metaphysical problems. Nonetheless, metaphysics is not the main goal of my inquiry, and I will not be perturbed if useful metaphysical upshots become unlikely.

I believe that it is ambition enough to spell out a constructive way to give meaning to the logical vocabulary and to investigate which logical laws are thereby underwritten. The metaphysical *consequences*, then, will not be pursued with much zeal, if at all.

On the other hand, if an argument requires strong realistic *assumptions* (such as Prawitz's appeal to an abstract realm of self-subsisting proofs that we shall meet in Sect.3.8), then I will point those out and proceed under the presumption that a Dummettian constructivist would shy away from making such assumptions.

Closely tied to this policy is something I mentioned above: When there will be cases in which correct assertibility and truth drift apart, I will keep my main focus on correct assertibility. Oftentimes, it will in such cases be quite unclear whether an acceptable notion of truth is available at all; I will often do not much more than flag

---

[30] Beall and Restall offer yet a third way to specify what a case is. A case might be a *situation*, something which might be incomplete and even inconsistent. This will give them a *relevant* logic, First Degree Entailment (FDE). I will say more about relevance and FDE in Chap. 3.

[31] If any. In fact, Beall and Restall argue that anti-realists can make use of classical logic (Beall and Restall 2006, p. 46 ff); their argument, however, has nothing much to do with their pluralism.

this worry and keep my attention on correct assertibility, whereas a metaphysician would undoubtedly not rest until the question whether a distinct notion of truth is available. After all, it seems clear that when truth and correct assertibility do not come to the same thing, any metaphysical conclusions should be drawn from the former, not the latter.

2.13 Chapter Summary

In this chapter, I have given a short survey of Dummett's expansive philosophical program. I tried to make the connections visible (and at least somewhat plausible) that Dummett sees between the analysis of language, the theory of logical consequence, and the most basic and important metaphysical questions. I showed how, based on a Wittgensteinian conception of meaning as determined by language use, Dummett thinks that certain semantic theories are unsuitable. We have seen that there is considerable uncertainty whether Dummett would prefer to think of meanings as conditions for correct assertibility or simply truth. If it is the latter, he would insist that the notion of truth that needs be employed has to be epistemically constrained to meet his conditions of manifestability and acquisition; hence, we are led toward a constructivistic theory of meaning. And such a constructivistic theory will, if thought through to the end, result in logical revision, or at the very least in a (metaphysically unexciting) form of eclecticism. As for a particular logic that we should embrace, Dummett campaigns for intuitionistic logic, which we will investigate in a more systematic manner in the next chapter.

Chapter 3
Intuitionism

3.1 Chapter Overview

As promised, this chapter should supply enough information about intuitionistic mathematics and its logic to make it possible to follow the rest of the book. I will begin by saying more about Brouwer's philosophy and the practical consequences for mathematics he and his followers drew from it. Next, I will present the logic and some semantical theories for it, with an emphasis on the BHK interpretation and the Kripke semantics.

3.2 Brouwer's Philosophy of Mathematics

Even if he might be one of its most influential propagators, intuitionistic logic is not Dummett's invention. It goes back to the Dutch mathematician L. E. J. Brouwer, who had in the early years of the last century developed a quite radical philosophy of mathematics. This philosophy was informed both by Kantian ideas and by Indian philosophy, and his writings are both fascinating and dense.[1] The main tenet, as I mentioned before, is that mathematical statements do not record facts in an abstract realm, but rather facts about mental constructions.

He developed this basic thought into a complex revision of classical mathematics. Many proofs that mathematicians accepted were ultimately based on a mistaken picture of the nature of reality. Unlike Dummett, who is driven by the need to explain language and logic, Brouwer was quite skeptical of the possibility of the correct communication of mathematical results. Formal logic especially could not hope to

[1] One should, for full effect, turn to his collected works. Especially, references to Indian mysticism have been purged in the reprints one finds in important anthologies such as Benaceraff and Putnam's *The Philosophy of Mathematics*.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_3,

capture all forms of valid mathematical reasoning.[2] Due to this disdain for formal logic, Brouwer never bothered to revise classical logic systematically. It was Brouwer's former teacher, G. Mannoury, who finally brought the project under way by promising a prize to whoever could axiomatize Brouwer's idea, and it was Brouwer's student, A. Heyting, who went on to win the prize by inventing intuitionistic logic.

3.3 Constructive Mathematics

Here is an example for the difference between classical and intuitionistic mathematics: Suppose we are wondering whether the equation

$$x^y = z$$

has solutions in which x and y are both irrational, while z is rational. Now consider the number $\sqrt{2}^{\sqrt{2}}$. If this number is rational, then we have an affirmative answer to our question. If it isn't, however, we can proceed as follows: Take x to be $\sqrt{2}^{\sqrt{2}}$ and set y to $\sqrt{2}$. Then, z is $\sqrt{2}^{\sqrt{2}^{\sqrt{2}}} = \sqrt{2}^{\sqrt{2}\times\sqrt{2}} = \sqrt{2}^{2} = 2$. Thus, whether or not $\sqrt{2}^{\sqrt{2}}$ is rational, we can answer our question affirmatively.

This is a proof that the classical mathematicians would acknowledge as valid. The intuitionist, on the other hand, will not be satisfied. For her, it is essential to answer which of the two cases hold. Is $\sqrt{2}^{\sqrt{2}}$ rational or not? Without an answer to this question the proof will, in her view, not succeed, because the intuitionist wants to see a *particular* number z that fulfills the requirement.[3]

In general, intuitionists will not accept a disjunction if they cannot discern which disjunct is provable. Under the realistic conception (which makes no sense to the intuitionist), one or the other has to be the case, whether we can come to know it or not. For the intuitionist, on the other hand, there is no similar guarantee that we will be able one day to effect the mental construction that is needed to decide the issue.

Here, then, is already one general feature that intuitionistic theories will have to satisfy: A disjunction will only be a theorem if either of the disjuncts is a theorem. This property is, naturally enough, called the *disjunctive property*, or *primeness* (a bit of algebraic jargon).

---

[2] An excellent resource for comparisons between Brouwer's, Heyting's and Dummett's views is Placek (Placek 1999). A shorter overview is in Chap. 7 of Shapiro (2000).

[3] Even though this example is repeated over and over again, it is a very recent trend to add the solution to the intuitionist's challenge: $\sqrt{2}^{\sqrt{2}}$ is irrational. The fact of the matter is that this has only recently been established, and the proof is said to be rather complicated (cf. Burgess (2009), p. 121).

3.4 The Axiomatization of Intuitionistic Logic

Building on such constraints as primeness and the failure of the Law of Excluded Middle,[4] Heyting gave his axiomatization of intuitionistic logic by expelling any axioms that struck him as suspicious from an axiomatization of classical logic. It is interesting and astounding that the BHK interpretation that we will see presently came quite a while *after* the axiomatization. One cannot but wonder how Heyting managed to decide whether to allow an axiom or not; the answer seems to be, fittingly enough, that he was guided by intuition only. In any case, here is the list of axioms he gave for propositional logic:

1. $A \supset (A \wedge A)$
2. $(A \wedge B) \supset (B \wedge A)$
3. $(A \supset B) \supset ((A \wedge C) \supset (B \wedge C))$
4. $((A \supset B) \wedge (B \supset C)) \supset (A \supset C)$
5. $B \supset (A \supset B)$
6. $(A \wedge (A \supset B)) \supset B$
7. $A \supset (A \vee B)$
8. $(A \vee B) \supset (B \vee A)$
9. $((A \supset C) \wedge (B \supset C)) \supset ((A \vee B) \supset C)$
10. $\sim A \supset (A \supset B)$
11. $((A \supset B) \wedge (A \supset \sim B)) \supset \sim A$

The only rule of inference is modus ponens. Adding either $A \vee \sim A$ or $\sim\sim A \supset A$ to this list will result in a full axiomatization of classical logic, as would the addition of further rules, such as classical reductio ($\sim A \supset B, \sim A \supset \sim B \vdash A$).

As I said, this logic was later supplied with a semantic theory that was thought to justify it in accordance with the concerns of constructive mathematicians. But also later developments in proof theory suggested that Heyting was on to something. Both of the Gentzenian revolutions in proof theory, the sequent calculus and the natural deduction calculus, accord a special role to intuitionistic logic. Indeed, an important part of Dummett's case rests on the argument that the natural deduction rules for intuitionistic logic are more harmonious (a technical term of his) and lead to more natural proofs. However, as I mentioned on p. 19, I will not go into these proof-theoretical arguments in this book.

3.5 Various Semantics

Over the years, there have been a number of semantical theories proposed for intuitionistic logic, some of which are[5]:

---

[4] Warning: In Brouwer's writing, this is called "Law of Excluded Third"; I follow Dummett in attaching this label to the double negation of LEM.

[5] See Dummett (2000).

1. The so-called Brouwer–Heyting–Kolmogorov interpretation (BHK for short), which spells out the meanings of the logical constants directly in terms of proofs.
2. The topological interpretation that interprets intuitionistic formulas as open sets in a topology.
3. Beth trees, in which different stages of mathematical investigation are related to each other, quite similar to the next item.
4. Kripke semantics, which I will introduce in detail below.
5. Kleene's realizability semantics, where formulas are interpreted as codes for algorithms.
6. Dialogue semantics, where intuitionistic validity comes to having a winning strategy in a dialogue game.

These interpretations are not the only ones around, but they already display the variety of guises a semantic theory might take, and their quite different strengths and weaknesses. An important spectrum along which they differ is how much they add to our intuitive understanding of the formulas versus how simple they make technical investigation. This distinction has been called "depraved versus formal."

At the extremes of this spectrum, we find the topological interpretation on the formal side and the BHK interpretation on the depraved. The topological interpretation tells us how we can translate our knowledge about topology to gain insights into the formal properties of intuitionistic logic; at the time, it was found, this meant a considerable information gain, because the topology of open sets was widely studied, while intuitionistic logic was a new topic of investigation. However, it is impossible to glean anything of the philosophical motivation from the association with open sets, and Dummett cannot see the topological interpretation deliver all he is hoping for in a semantics:

> In attempting to find a semantics with respect to which a given logical system is both sound and complete, a logician is not merely seeking an algebraic rather than a proof-theoretic characterization of the deducibility-relation of that system: a semantical theory proper (...) is to be distinguished from a merely algebraic valuation system (such as the topological interpretation (...)). If any distinction between a semantical theory and a purely algebraic valuation system can be drawn at all, the ground of it must lie in the fact that the semantical theory is connected with the way in which both logical and non-logical expressions are given meaning, while a purely algebraic one is not (Dummett (2000), p. 256).[6]

3.6 The BHK Interpretation

In contrast to the topological interpretation, the BHK interpretation does give us the meaning of the logical constants and is thus a candidate for a semantic theory in Dummett's sense. These meanings are spelled out in terms of *proof conditions*.

The BHK interpretation gives us a very good idea of what it is to prove a complex statement, provided we know what it is to prove statements of lesser complexity.

---

[6] See also TOE, p. 293, and Dummett (1998), p. 126.

To get a basic idea of how this will work, consider a conjunction of two statements, the proof conditions of which we assume to know. That is, we could come to recognize a proof of either of them if we were presented with one. What, now, would a proof of the conjunction look like? Well, it would simply be a proof of the first conjunct, followed by a proof of the second conjunct.

The case of disjunction is equally trivial, but what about negations? The intuitionist tries to express everything in terms of constructions and proofs that are obtainable by us. A negation, in contrast, seems to be telling us something about the *impossibility* of proof and construction. How might this be expressed in constructivist terms?

Consider how we go about proving negative statements in mathematics, such as the claim that there is no greatest prime number. The first step is to assume that there is a greatest prime number. Then, we reason from that assumption, until we hit a contradiction. In the case of Euclid's famous proof, we prove that there is a greater prime number than the one we assumed to be the greatest, which obviously contradicts the assumption.

The proof of a negation is thus a fairly elaborate process, and it seems at heart to be a kind of conditional. *If* there is a greatest prime number x, *then* there is another one that is even greater than x. Or, to make the contradiction perfectly explicit: *If* there is a greatest prime number x, *then* x is the greatest prime number and x is not the greatest prime number.

This leaves us with two questions: How does the constructivist explain *conditionals*? We will see that below by directly considering the BHK clause, but before that, let us consider the second question: Isn't our definition of negation circular? We are trying to explain what negation means by taking recourse to a conditional that has a contradiction as its consequent. But a contradiction is obviously a notion that presupposes the notion of negation.

The intuitionists employ the following trick: Instead of a contradiction, they use a mathematical statement that can be assumed to be false, usually $1=0$. This absurd statement is abbreviated as $\perp$. Mathematics is conceived as endowed with such an inherent coherence that the assumption of *any* false statement will eventually lead to a proof of $1=0$.

Here is one version of the clauses for the sentence connectives:

- c is a proof of $A \wedge B$ iff c is a pair $(c1, c2)$ such that $c1$ is a proof of A and $c2$ is a proof of B
- c is a proof of $A \vee B$ iff c is a pair $(c1, c2)$ such that $c1$ is a proof of A or $c2$ is a proof of B
- c is a proof of $A \supset B$ iff c is a construction that converts each proof d of A into a proof $c(d)$ of B
- nothing is a proof of $\perp$
- c is a proof of $\sim A$ iff c is a construction which transforms each proof d of A into a proof $c(d)$ of $\perp$.

We now see how conditionals are dealt with. A conditional is taken to be a construction that delivers a proof of the consequent if supplied with a proof of the antecedent.

We also see what was noted above, that the clause for negation has the same structure as the clause for the conditional.

Indeed, often enough, the last clause is left out, because negation is simply *defined* as an implication of absurdity: $\sim A$ is short for $A \supset \bot$. Again, the idea that is supposed to be captured is that, when you are able to derive an absurdity from a hypothesis, you can be sure that you will never come across a proof for that hypothesis.

It is the explanation of complex statements' semantic values as something inherently recognizable (proof conditions) that makes the intuitionistic account so attractive to Dummett:

> The intuitionistic explanations of the logical constants provide a prototype for a theory of meaning in which truth and falsity are not the central notions. The fundamental idea is that a grasp of the meaning of a mathematical statement consists, not in a knowledge of what has to be the case, independently of our means of knowing whether it is so, for the statement to be true, but in an ability to recognize, for any mathematical construction, whether or not it constitutes a proof of the statement; an assertion of such a statement is to be construed, not as a claim that it is true, but as a claim that a proof of it exists or can be constructed. The understanding of any mathematical expression consists in a knowledge of the way in which it contributes to determining what is to count as a proof of any statement in which it occurs. In this way, a grasp of the meaning of a mathematical sentence or expression is guaranteed to be something which is fully displayed in a mastery of the use of mathematical language, for it is directly connected with that practice. (...) [O]ur understanding of a statement consists in a capacity, not necessarily to find a proof, but only to recognize one when found (WTM p. 70).[7]

To get a full semantic theory, an explanation of the meaning of *all* statements would be needed. However, the BHK interpretation provides at most what Dummett calls a "skeletal" semantics. It gives an account of complex statements, but tells us little about atomic ones. If we restrict our attention to arithmetic, as Dummett often does in his examples, the atomic statements can be seen as numerical equations and proofs for them as basic arithmetical computations. However, if we want to extend the account to all of mathematics, it is not at all clear what the best notion of a proof for an atomic statement would be.

---

[7] This seems to be one of the points at which a certain amount of idealization of our capacities is required. It would not seem right to attribute understanding of Fermat's Last Theorem only to those few who could recognize Wiles's proof to be correct. Even worse is the proof of the Four Color Theorem, which has been proven, but only with the aid of computers that surveyed an extensive range of cases too large for any single human being to check. Rather than to say that we really do not understand these mathematical theorems, N. Tennant proposes the following refinement of Dummett's requirement (picking up an idea he attributes to J. Cogburn):
"For a speaker S to be credited with a grasp of the meaning of a sentence ϕ we should have good grounds for believing that, if presented with some finite piece of discourse π, S would be able to deliver a correct verdict on any aspect of π that is relevant to arriving at a correct judgement of the form 'π is a proof of ϕ' or of the form 'π is a disproof of ϕ' or of the form 'π is neither a proof nor a disproof of ϕ', that is, for any such aspect α, S would, after some time, be able to judge whether α was as it ought to be, in order for π to have the status in question (Tennant (2002), p. 154)."
In other words, one has to be able to recognize each step in the proof as correct or incorrect, even if one cannot survey the whole proof.

However, a skeletal semantics by itself is in principle enough to specify logical consequence. To say that a statement follows from a given set of premises is to say that, if the premises are provable, the consequence is provable as well. If the statement follows from the empty set of premises, its provability is based on nothing but its logical form and the way the logical constants are explained.

The LEM would, under the BHK interpretation, be saying that it is possible to give an intuitionistic proof for every given statement or its negation. But we actually have no guarantee for such a claim. The only such guarantee that an intuitionist would accept would be a proof (or at least a secure method of obtaining a proof) of the statement in question or of its negation.

That the LEM should appear dubious to a verificationist was clear even before we saw this concrete interpretation. More interestingly, under the BHK interpretation, we can now also begin to understand why Double Negation Elimination should fail. Replacing the negations by implications of absurdity, the law can be formally stated as: $((A \supset \bot) \supset \bot) \supset A$. This would, were it valid, tell us the following (which I try to make easier to parse by employing italics, emphasis and line breaks):

> If we have a CONSTRUCTION THAT TURNS
>
> a *further construction, which turns a proof of A into a proof of* $\bot$
>
> INTO A PROOF OF $\bot$,
>
> then we could always transform this whole thing into a proof of A.

In other words, if we can be sure that we will never find a proof for the claim that we can be sure never to find a proof of A, then this is as good as having a proof of A. But the intuitionist holds that it is not: It serves only to show that $\sim A$ will never be proven. We might well know that, without yet having a guarantee that we will find an actual constructive proof for A.

It will be good to see an example here. As I mentioned, the Law of Excluded Third, $\sim\sim(A \vee \sim A)$, is provable in intuitionistic logic. This tells us that DNE cannot be valid, else we would get a proof of LEM. I will show how the Law of Excluded Third follows from the BHK clauses.

A proof of $\sim\sim(A \vee \sim A)$ is a method of transforming a proof of $\sim(A \vee \sim A)$ into a proof of an absurdity, $\bot$.

Let us assume, then, that we have a proof of $\sim(A \vee \sim A)$, i.e., a method of transforming any proof of $A \vee \sim A$ into a proof of $\bot$. Such a proof might either be a proof of A, or it might be a proof of $\sim A$.

So, to be sure that we can transform any proof of $A \vee \sim A$ into a proof of $\bot$, we need a method to transform a proof of A into a proof of $\bot$ (as well as a method to transform a proof of $\sim A$ into a proof of $\bot$).

Under the assumption that we have a proof of $\sim(A \vee \sim A)$, then, we see that we have a method to transform a proof of A into a proof of $\bot$. That is, we have a proof of $\sim A$.

A proof of $\sim A$ in turn amounts to a proof of $A \vee \sim A$. Under our assumption, we have a method of turning this proof into a proof of $\bot$.

That is, given a proof of $\sim(A \vee \sim A)$, we can now give a method of deriving a proof of $\bot$. And this completes our proof of $\sim\sim(A \vee \sim A)$.

Intuitively, the validity of the LET tells us that we can be sure that, even though we cannot assume the validity of LEM, we know that we will never find a *counterexample* to it. This is why it seems that we cannot express the constructivist objection to LEM in terms of absolutely undecidable propositions. The temptation is to express their worry by saying that neither Goldbach's conjecture nor its negation might be provable. But that seems to be saying that it is a possibility that a negation of an instance of LEM is true, something we have just seen is not an option for the intuitionist.

It is debatable whether a stable philosophical position can be built around this apparent dilemma.[8] The intuitionist will have to say that there are no absolutely undecidable statements, but that there still is no guarantee that, given a *pro tempore* undecidable statement, we will find a way of deciding it.

3.6.1 Correctness and Explosion

Let us come back to the BHK interpretation. We have seen that we can deduce some characteristic features of intuitionistic logic from this interpretation. In general, however, it is doubtful whether it is concrete enough to support correctness (completeness and soundness) proofs for intuitionistic logic, even if S. Buss writes:

> It is not difficult to see that the intuitionistic logic (...) is sound under the BHK interpretation, in that any provable formula has a proof in the sense of the BHK interpretation (Buss 1998, p. 65).

Historically, people have had more problems seeing this than Buss allows. As I mentioned above, the BHK interpretation takes its name from Brouwer, Heyting, and Kolmogorov. Other than giving the initial impetus to intuitionism, Brouwer's contribution to the BHK interpretation is quite negligible, and the inclusion of his name seems little more than an honorary act.

Heyting and Kolmogorov, on the other hand, had in fact both independently proposed BHK-style interpretations. However, they were not at all in agreement whether intuitionistic logic is sound with respect to their interpretation. While Heyting did indeed think so, Kolmogorov was worried[9] about the admissibility of the intuitionistically valid inference known as *ex contradictione quodlibet* or, as I will mostly call it, Explosion. It allows the intuitionist to draw an inference from a contradiction to an arbitrary statement ($A \wedge \sim A \vDash B$). The question is whether there is a general method of generating a proof of any statement whatsoever from a proof of a statement and a proof of its negation. As $\sim A$ can be seen as the conditional $A \supset \bot$, A and $\sim A$ together let us infer, via modus ponens, $\bot$. To argue for ECQ, we have to give an argument that $\bot$ entails every other statement.

There are two conceivable ways for the intuitionist to argue for ECQ.

---

[8] See Williamson (1994), who adds insult to injury by arguing that the intuitionist cannot even say that, by way of historical coincidence, there is a chance that we might not actually decide whether Goldbach's conjecture or its negation holds.

[9] Cf. van Dalen (2004).

The first way of doing this is to say that it will indeed be possible to find a way to transform a given proof of $1=0$ into a proof of any other statement. This is clearly constructive in spirit, if indeed it can be pulled off. However, it is not quite clear that it is a plausible assumption if we consider all of mathematics, not just elementary arithmetic or some other restricted area. An additional worry is that we seem here to be transforming proofs that can not exist[10]; it isn't perfectly clear how we can claim to be able to do that.

The second way is to say that it is quite all right to claim to have a construction that will effect the conversion of a proof of the premise into a proof of the conclusion, simply because there will be no input that would put our claim to the test.

Here, the problem is that it is not wholly clear that the argument is all that constructive. It plays much more on the impossibility of constructions than on their possibility (or their possibility in principle). We claim to have a construction where in fact we do not have one, or one that ostensibly would not do the required job if, counterfactually, it were supplied with a suitable input.

Intuitionists have wavered between these two ways of motivating ECQ, and I will come back to the distinction when I discuss negation in empirical contexts. As they will come up again, let us find two names for these lines: The first might be called a *truly constructive conversion*, the second an *empty promise*[11] *conversion*.

Kolmogorov and others after him thought that there is something fishy about both ways this argument could be played out; for example, worries about this argument have led to the development of *minimal logic*, where a contradiction does not entail everything. Minimal logic[12] is axiomatized by the axioms of intuitionistic logic except axiom 10. This axiom, $\sim A \supset (A \supset B)$, is what makes ECQ hold in intuitionistic logic. The notions minimal logicians want to capture are not distinct from the ideas encapsulated in the BHK interpretation, so Buss's claim that intuitionistic logic is obviously sound with respect to it seems questionable.

We are then looking for something more to the center of the spread between the formal extreme of the topological interpretation and the depraved BHK interpretation. What I will be concentrating on from now on are the Kripke semantics for intuitionistic logic, both because they strike a good balance between formality and intuitiveness and because they are so widely employed that incorporating ideas from other logics, as I will do in the later parts of the book, is particularly easy.

3.7 Kripke Semantics

The semantics is a variation in the well known possible worlds semantics of modal logics. The worlds are supposed to be information states that the inquisitive subject steps through in her quest of knowledge. As intuitionistic logic was conceived of as

[10] Cf. Wansing (1993), p. 21.

[11] "I'll give a million dollars to the local orphanage... when pigs fly."

[12] Introduced in Johansson (1936).

a logic for constructive mathematics, the worlds are normally taken to record what has been proven up to a certain point in time. There is an accessibility relation on the worlds that has a strong temporal flavor: One stage comes literally *after* the other. Let us spell this out a bit more formally:

A model will be a structure, $[W, \leq, v]$, where W is a non-empty set of worlds or information states, and $\leq$ is a binary relation on those worlds which is reflexive, transitive, and anti-symmetric, that is, a partial order.

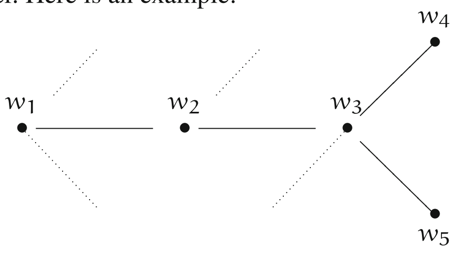
The valuation function v assigns a truth value, 1 or 0, to each atomic statement p at each world. Intuitively, for p to receive value 1 at a world means that that statement is proven at that world. For it to receive value 0, on the other hand, means just that p has not been proven (yet) at world w. It does *not* mean that p has been disproven at world w. That is, if the investigation is carried on, we might well find a proof of p at a later stage or world, and thus the value at that stage will turn to 1. On the other hand, a proof is something definitive, such that a statement that is proven at a world will always remain proven at subsequent stages.

In the semantics, this thought is manifested in the so-called *heredity constraint*:

For each p: if $w \leq w'$ and $v_w(p) = 1$ then $v_{w'}(p) = 1$.

For $v_w(p) = 1$ I will write $w \Vdash_1^v p$, and for $v_w(p) = 0$ I will write $w \Vdash_0^v p$, although I will often leave out the superscript v to reduce clutter. I will also write $w \nVdash_1 p$ for "not $w \Vdash_1 p$," and likewise $w \nVdash_0 p$.[13] As the valuation functions are total, $w \Vdash_0 p$ is equivalent to $w \nVdash_1 p$.

One can represent the models graphically, which makes understanding what is going on much easier. Here is an example:



	w_1	w_2	w_3	w_4	w_5
p	0	1	1	1	1
q	0	0	0	1	1
r	1	1	1	1	1
s	0	0	0	1	0

This shows part of a model for intuitionistic logic. The lines represent the accessibility relation between the worlds, represented as dots and labeled w_n. It is meant to be read from left to right, such that, for example, w_1 should be seen as accessing w_2, but

[13] Also, expect to see $t \geq s$ as a notational variation of $s \leq t$ occasionally.

not vice versa. Because the relation is by definition reflexive, every world can access itself, a fact that is not represented in the diagram. If there are self-accessible worlds and others that are not self-accessible, it makes sense to indicate the self-accessing ones; one will typically see such worlds represented thus:

w_n

As *all* the models we will meet in this work will be reflexive, these loops will be left out to reduce clutter, but they should be understood implicitly.

The dotted lines indicate other branches in the model that we choose not to consider. Note especially the dotted line leading to w_3: These models need not be trees, and a world may have more than one predecessor.[14] Remember that the worlds are supposed to record information states; the idea that lies behind allowing worlds with more than one predecessor is that one may come to the same state of information via different routes. Say we will from now on turn to prove two mathematical statement, say, the Goldbach conjecture and the P = NP problem. One way to do this is to first turn to Goldbach, and then take care of P = NP, another way is to do it the other way around. If we solve nothing but those two problems, we should come out at the same state of information, no matter in which order we tackle them.

Now, take a look at the valuation below the diagram. In this matrix, the fate of four propositional constants is recorded. The workings of the heredity constraint come out clearly here: Once a proposition is proven, that is, once it receives value 1, it stays proven. At the first world, the only thing we have proven is the proposition r, and will never go back to being unproven (receiving value 0), no matter what else happens. At w_2, we manage to prove p, and so we have two atomic propositions that get assigned value 1.

Passing from w_2 to w_3 seems to bring little change. There might be other statements that we have not listed, but maybe there aren't. There is nothing precluding two or more successive worlds to be indistinguishable in terms of the statements that are proven at them. If we want to think of the worlds as marking points in time, then in such a stretch of indistinguishable worlds, there is simply no progress made in our inquiry.

In contrast, we register an impressive increase in our knowledge when we move from w_3 to w_4: Both of the remaining unproven atomic statements, q and s, are proven simultaneously. Again, there is nothing in our definition of intuitionistic models preventing this.

The transition from w_3 to w_5 marks a different possible way our investigations may unfold: Only one of the statements is proven, while the other stays open.

[14] As far as intuitionistic logic is considered, it does not *hurt* to restrict one's attention to trees. The consequence relation that arises from tree models is not different from the consequence relation one gets from the more inclusive set of partially ordered models. This changes when co-implications are added to the vocabulary of intuitionistic logic, cf. Restall (1997). But I shall leave co-implications for a later point in the story.

3.7.1 Logical Constants

So much for atomic statements. Next, here are the conditions that determine the semantic values of logically complex statements:

For all $w \in W$:
$w \Vdash_1 A \wedge B$ iff $w \Vdash_1 A$ and $w \Vdash_1 B$
$w \Vdash_1 A \vee B$ iff $w \Vdash_1 A$ or $w \Vdash_1 B$
$w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \Vdash_0 A$ or $x \Vdash_1 B$
$w \Vdash_1 \sim A$ iff for all $x \geq w$, $x \Vdash_0 A$

Under these conditions, the heredity constraint holds for all statements, not just the atomic ones.

I think that the success of the semantics lies in the fact that most people see these clauses as a formal precisification of the BHK interpretation, and not really a rival semantic proposal. Let us see in how far the two accounts of complex statements are in harmony.

To repeat, the inquiring subject knows a statement to be provable at a given world if she has a proof of it (I will say more about what exactly it should be taken to be necessary and sufficient to "have" a proof in 3.8). If that statement is a conjunction, then that proof consists of a pair of proofs, one for each conjunct. Therefore, she has these proofs at her disposal, and thus, the two conjuncts are proven on their own. That is, at the world in question, each of the conjuncts will get the value 1 assigned to it, just as the clause above tells us.

Disjunctions are just as straightforward. If our subject has a proof of a disjunction, then BHK tells us that she has something that amounts to a proof of either disjunct. Therefore, either the first or the second disjunct will get the value 1 at the world we are concerned with.

It gets interesting when we look at the intensional notions, that is, those that look to other worlds in order to determine the truth value of a statement at the present world. A glance back at the clauses tells us that this concerns conditionals and negations.

First of all, the intensional notions make clear that knowing whether a statement receives value 1 or value 0 at a world is not enough to know the semantic value of the statement, in the sense of Dummett. The semantic value is not just telling us whether a statement is assertible or not, but also how a complex statement in which the statement is a constituent should be evaluated. For this, we will need to know the fate of the statement in the accessible worlds, and that must then be seen as part of the semantic value as well. It is easy to overlook this, due in part to somewhat unfortunate terminology, but also due to the fact that one tends to focus on the value 1 too much. Given the heredity constraint, it makes no difference whether we say that the semantic value of a statement that receives value 1 is exhausted by that, or whether we say that we need to know the future development of that statement, for in all accessible worlds it will also receive value 1. If the value is 0, on the other hand, we cannot say whether, for example, the negation of the statement receives value 1 or value 0. This will depend on the question whether the original statement will receive value 0 in all subsequent world or not.

With that in mind, let us look at how exactly these intensional notions work. A proof of a conditional, BHK tells us, is a construction that takes any proof of the antecedent and turns it into a proof of the consequent. That is, if a conditional is true at a world, it is inconceivable that the inquiring subject should acquire a proof of the antecedent and not be able to prove the consequent. Therefore, we see that in all possible developments the investigation might take from the point in which the conditional is proved, a proof of the antecedent will always immediately result in a proof of the consequent: In all worlds that represent possible developments, either the antecedent is unproven (has value 0) or the consequent is proven (has value 1).

A negation is proven, the BHK interpretation explained, if we have a construction that will turn any proof of the negated statement into a proof of an absurdity. Such an absurdity, of course, will never be proven, and thus, there will be no world at which it will receive value 1. Therefore, no world that is a possible development of the present stage of investigation will assign value 1 to the negated statement, or we would be forced to assign 1 to the absurdity.[15]

This tells us something about the kind of modality that is involved in the Kripke interpretation. The possible developments the interpretation talks about can only be epistemically possible, not mathematically: If they had to be mathematically possible, then no world would assign value 1 to a false mathematical statement, and thus the negations of all false statements would be true from the outset.

3.7.2 Logical Consequence

You will remember from the last chapter that once a semantic theory has informed us about the semantic values of atomic statements and how to compute the values of complex statements on that basis, it will have to tell us how these values relate to truth. For one thing, because we are interested in truth per se, but more importantly because we want to define logical consequence in terms of truth preservation.

The semantic values in our Kripke semantics are 1s and 0s. I said that a statement receives value 1 at a stage or world iff it is proven at that stage. Of course, once this happens, it is clear that the statement is true. However, how about those stages that came before, where the now proven proposition had not yet been proven? Clearly, the semantics tells us that it will have received value 0 at these stages. But do we want to say that it had not been *true* at those earlier stages, even though we went on to prove it?

This is a very tough question for the constructivist, and I will come back to the question of how the semantic values and truth relate at the end of this chapter.

---

[15] That the BHK interpretation is alluded to in this paragraph is not accidental to the exposition. I will argue presently (in Sect. 3.7.3) that BHK and Kripke semantics need each other to satisfy Dummett's needs.

For now, let us assume that in our models, only those statements are true that receive value 1, and that this value is what the consequence relation has to transmit:

$\Gamma \vDash A$ iff in every model and every $w \in W$, if $w \vDash_1 B$ for every $B \in \Gamma$, then $w \vDash_1 A$.

With consequence thus defined, the Kripke semantics is sound and complete with respect to Heyting's axiomatization. In particular, we can see that *ex contradictione quodlibet* will hold, because a contradiction is false at every world in every model.

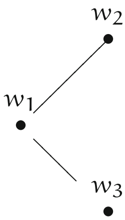
It is not hard to come up with counter models for the characteristic failures of classical validities in intuitionistic logic:

Let $w_1 \Vdash_0 p$ and $w_2 \Vdash_1 p$. This is a counter model for both LEM and DNE.

LEM: We have that $w_1 \Vdash_0 {\sim}p$ because p does not receive value 0 at all subsequent worlds. Therefore, at w_1, $p \vee {\sim}p$ receives value 0.

DNE: We obviously also have that $w_2 \Vdash_0 {\sim}p$. Thus, ${\sim}p$ receives value 0 at all worlds accessible from w_1. But that means that $w_1 \Vdash_1 {\sim\sim}p$, and from that we can compute that $w_1 \Vdash_0 {\sim\sim}p \supset p$.

Here is a counter model to the one De Morgan's law that fails in intuitionistic logic, ${\sim}\ (p \wedge q) \supset{\sim}\ p \vee {\sim}\ q$ (all the others hold):



Let $w_2 \Vdash_1 p$, and $w_3 \Vdash_1 q$, with no further assignments of value 1 to either p or q. For all $k \in \{1, 2, 3\}$, $w_k \Vdash_0 p \wedge q$, therefore $w_1 \Vdash_1 {\sim}\ (p \wedge q)$. However, $w_1 \Vdash_0 {\sim}\ p$, because there is an accessible world, w_2, for which $w_2 \Vdash_1 p$. Likewise, $w_1 \Vdash_0 {\sim}\ q$. Thus, $w_1 \Vdash_0 {\sim}\ p \vee {\sim}\ q$, and that tells us that $w_1 \Vdash_0 {\sim}\ (p \wedge q) \supset{\sim}\ p \vee {\sim}\ q$.

3.7.3 Kripke Semantics and the BHK Interpretation

Now, let us pause to ask the following: If we had nothing but the Kripke semantics in front of us, would it fare better than the topological interpretation? Is it clear what the clause for negation is expressing?

If we want to give more concrete content to the BHK interpretation, then the Kripke semantics work rather well; at least it gives one way of answering the questions that were not quite clear about BHK, for example whether *ex contradictione quodlibet* should hold or not. But do we even need the BHK interpretation, or would the Kripke semantics alone suffice?

The answer seems to be that we have a mutual dependence here: The Kripke semantics will need the heuristics of the BHK interpretation (or something like it) to put it in a better position than the topological interpretation.

For example, consider the clause for negation: A negation is true iff the statement that is negated is unproven in all accessible worlds. But what if we are in a model in which it just happens, for no particular reason at all, that in all accessible worlds from the actual world, there is a statement that is unproven throughout? The negation of that statement will be true, but why? Isn't this just chance?

Similar questions arise about the conditional. In a model in which there just happens to be no world in which the antecedent is proven and the consequent isn't, the corresponding conditional will be true throughout. But it could be quite a coincidence that matters stand as they do in the model. This would seem somewhat unsatisfying, just as the classical, material conditional is unsatisfying, as it holds between pairs of statements that might be quite unrelated to each other, as long as the first is true and the second false.

To begin to explain we should again mention, as I did above, that the alternative worlds have to be seen as epistemic alternatives. To make sense of negations and conditionals, we need to provide accessible worlds for all the epistemic alternatives that we can conceive of.

This, as it happens, puts some pressure on the intelligibility of small models like the one I gave above as counterexample to DNE. There were just two worlds, with the first just accessing the second and itself, and the second only itself. But what states of investigation are these worlds supposed to be? In the first, there is no epistemic possibility of p not being proven in the next step of the investigation. It is as if we had an algorithm for proving the statement and knew that, come what may, we will apply it. In such a case, it is actually tempting to consider the statement proven at the current world already.

These are not the only problems. In fact, all finite models seem problematic. In finite models, we have the difficulty that having a proof of $\sim\sim p$ means that p will be proven sooner or later, no matter where the investigation may lead. If not before, then in the end nodes, those that access only themselves, p will be proven. That is not very satisfying; we want to make room for the possibility of proving $\sim\sim p$ without any guarantee for finding a proof of p at all.

Finite trees seem to be rather at odds with the interpretation we are after, then. Another problem with finite models is that the behavior of end nodes is strange: As they access no worlds but themselves, we find that for all statements q that have not yet been proven, $\sim q$ will become proven. After all, the qs are unproven in all accessible worlds, and nothing more is asked for. This is another feature that makes finite models questionable, for it seems quite unintuitive that a lot of negations should become provable at the end of investigation.

Even if we resolve to consider only infinite models.[16] it seems that to get the Kripke semantics to confer anything like concrete meaning to the logical constants, one has to add some explication of why the models look like they do. Very well,

[16] Happily, this has no effect on the logic.

there might be no state of information that the current one might develop into that would prove p. So, the semantics claims, we know that we have a proof of $\sim p$.

But why *is* there no proof of p in the subsequent worlds? What makes us look at such a model, and not one in which there is a later proof of p? The answer is that the fact that there will be no such world is already contained in the information available at the actual world. There is something that we know that tells us that p will not be verified.

It seems that the models themselves, or at least what we have said about them until now, do not give us a good answer to the question what exactly it is that we know here. To give such an answer, we have to supply some additional explanation. If this explanation is to be true to the constructive spirit of intuitionistic logic, we presumably have to go back to an account in terms of proofs. Such an account is, in presentations of this semantics, often given in a more or less hand-waving manner. However, we have already seen a relatively elaborate account that we can utilize, namely the BHK interpretation.

Indeed, the BHK clause for negation gives us quite a good indication of how to model a state of information in which a negation is proven. It is the fact that we have a means of turning a proof of p into a proof of $\perp$ that lets us choose a world in a model in which no successor world proves p as a suitable representative of such a state of information. Likewise, the models in which a conditional is proven at a world look the way they do (i.e., the worlds the world can access will either assign value 0 to the antecedent or value 1 to the consequent) because we know of a method of transforming a proof of the antecedent into a proof of the consequent.

It seems then that we really need a BHK-style interpretation to supplement a Kripke semantics to give meaning to the logical constants of intuitionistic logic.[17] On the other hand, the Kripke semantics supply enough mathematical detail to give soundness and completeness proofs. For example, it will be clear from the clauses for the connectives and the definition of validity that *ex contradictione quodlibet* is valid under the BHK interpretation, as precisified by the Kripke semantics.

That is of course not quite a conclusive proof that the right way to understand the BHK interpretation is to say that it validates *ex contradictione quodlibet*. We could just as well have given a Kripke semantics for minimal logic, the logic that does not validate this inference.

Indeed, such a semantics would not look too different from the Kripke semantics we have just seen. The only difference would be that we drop the requirement that $\perp$ should always receive value 0. Then, a contradiction would not entail everything any more, for it might be assigned value 1 at a world, provided $\perp$ receives value 1 as well at that world.

Now, an absurdity constant that need not be false might seem to be a strange artifact that we would like to avoid. Thus, one could run the argument that, if Kripke semantics

[17] Dummett comes to a similar conclusion, although he is concentrating in his discussion on Beth trees (Dummett (2000), p. 287). The difference between Kripke models and Beth trees is not too great, at least in those respects that interest me here. I choose to discuss Kripke models because it is much easier to compare them to the semantics for the logics that follow (at least for those logics that have been characterized and given semantics in the literature).

are the best way to give a more concrete elaboration of the BHK interpretation, intuitionistic logic responds to our intuitions better than minimal logic. The minimal logicians might deny the strangeness of the fact that $\bot$ need not be false, or, maybe more plausibly, they might propose a different, non-Kripkean, semantics that fits their intuitions better.

In sum, the Kripke semantics for intuitionistic logic and the BHK interpretation do seem to make a good couple. Together, they explain how complex statements receive their semantic values from the values of their constituents.

3.8 Tensed or Untensed Provability?

Now, if we want to progress to a notion of logical consequence, we have to explain how these semantic values relate to correct assertibility or to truth, or else give a surrogate notion that is to be identified with correct assertibility or truth and explain how the semantic values relate to *that*.

The natural surrogate notion that falls out of our semantics is that of *provability*. There is a certain ambiguity hidden in the modality, though. Certainly, a statement that is proven is also provable. However, the modalization hints at the intent to include more statements than just those that are actually proven. Here are four proposals to spell out what the provability of A might come to:

1. We have a proof of A.
2. (1) or we know of a method of obtaining a proof of A.[18]
3. (2) or we know of a decision method for A (A is decidable), and that method would give us a proof of A if we applied it. However, for all we know, it might have given us a proof of $\sim A$ instead.
4. There *is* a proof of A, whether we know or ever will know of that proof or not.

Note that the fourth option does not make decidability a prerequisite for provability. As a consequence, this kind of provability is an untensed notion, whereas the first three options are tensed: A statement might *become* provable.

From the outset, none of these options seems patently absurd as an interpretation of "provable." However, we need to ask how well they fit to the notions we are ultimately aiming at, correct assertibility and truth.

If we look at truth, then options 1–3 make truth an essentially tensed notion as well. Statements become true as we gain more and more knowledge. The main problem with this, it seems, is that it simply goes against all intuitions about truth that we

[18] One may not wish to think of this option as anything distinct from option 1. If we know of a method of obtaining a proof, then we *have* a proof, or else we would not know that we would obtain a proof. This depends a bit on how closely individuated proofs are taken to be, and on how strict we are with the term "proof." For example, if an informal but valid argument is a proof, then there might be not much point in drawing the distinction between (1) and (2). Nothing much in the following argument will change if they are collapsed.

have. We feel that truth in general, and especially mathematical truth, is made for eternity.

On the other hand, the thought that a statement could fail to be correctly assertible at one time, but become assertible at a later time, is much less troublesome. Indeed, here the intuitions might well be reversed, the notion of eternal correct assertibility might strike one as artificial and strange. In that case, a tensed notion of provability would be quite welcome. I discuss the differences between the three tensed options below, first let us look at the untensed one, option 4.

3.8.1 Untensed Provability

The untensed provability of option 4, while it might cater to the intuition that truth is untensed, seems dangerously close to the metaphysical and semantic doctrines of realism that the anti-realist was hoping to avoid. It sounds as if it makes reference to a realm of proofs that need not at all be amenable to epistemic inspection.

D. Prawitz has argued that this is not a problem for an anti-realistically minded theoretician.[19] Because the realm in question is populated by *proofs*, it cannot contain anything that is capable of outstripping our recognitional capacities, at least when they are idealized in rather harmless ways.

But still, there are grave problems with such an account of provability: First, it is very hard to see how reference to such a realm could support doubts of bivalence. Indeed, Dummett makes exactly this point against Prawitz:

> There is a well-known difficulty about thinking of mathematical proofs - and, equally, of verifications of empirical statements - as existing independently of our hitting on them, while insisting that they are proofs we are capable of grasping or giving fails to resolve. Namely, it is hard to see how the equation of the falsity of a statement (the truth of its negation) with the non-existence of a proof or verification can be resisted: but, then it is equally hard to see how, on this conception of the existence of proofs, we can resist supposing that a proof of a given statement determinately either exists or fails to exist. We shall then have driven ourselves into a realist position, with a justification of bivalence (Dummett (1987), p. 285).

A constructivism such as Prawitz advocates has thus actually quite a lot of similarity with realism, as construed by Dummett, and is likely to be susceptible to the same kinds of complaints.[20]

Second, it is very unclear that Prawitz's notion of provability has a connection to our linguistic behavior that is any closer than that of the realist's epistemically completely unconstrained truth. To say something that is provable cannot, by itself, be sufficient to make an assertion that will be judged correct, for no one need be able to find out about the provability of the statement. If our main focus is on correct assertibility, then option 4 seems unattractive.

---

[19] Prawitz (1987), p. 157.

[20] See Martino and Usberti (1994) for an extended argument against Prawitz's proposal and untensed constructive truth in general.

Lastly, there is a practical problem: It is not at all clear how we could connect our Kripke semantics to Prawitz's idea. The value 1 is surely not to be equated with tenseless provability or truth. It makes little sense to say that something has always been provable at one stage, even though it might not have always been provable before.

We might try to say that the actual path our investigation is going to take will tell us what is provable. But even if such a "thin red line" were an object that a constructivist could employ (and this is obviously a questionable proposition), it still would not be guaranteed to give us all that is provable. We might well not hit one of Prawitz's proofs, no matter how long we go after it.

But then it seems we cannot capture the notion of tenseless provability with our Kripke models, and consequently, the logic we can characterize by them is not at all sure to correspond to the logic of provability. Whatever might be a semantic theory that embraces Prawitz's untensed provability, this is not it, at least without some essential augmentation.

3.8.2 Tensed Provability

Let us then direct our attention to the tensed alternatives 1–3.

As I said above, the main worry of allowing a tensed notion of *truth* seems to be that it fails to match our intuitions about truth.[21] Truth, we would say, is nothing that comes and goes. Maybe we could get over this when it comes to future tensed statements: It seems quite plausible to hold that a statement about tomorrow's weather is neither true nor false, but becomes true or false tomorrow. But statements about mathematics? Can we really make peace with the idea that Fermat's Last Theorem is true today, but was not true in 1993?

---

[21] Dummett notes that there is another property usually ascribed to truth that we would have to do without:

> It is worth noting that beside the tenselessness, another feature that is commonly assumed to hold for truth seems to fail for the notion of truth at hand: This notion of truth does not satisfy the disquotational scheme, at least if negation is understood in the intuitionistic way (LBM p. 166).

From

"p" is true iff p

it follows intuitionistically that

"p" is not true iff $\sim p$

but the "only if" part is wrong: For "p" to be not true, it is enough that we are in a situation where p is undecidable. More doubts about the unrestricted validity of the disquotational scheme can be found in Dummett (2004), p. 14 ff.

Of course, the constructivists might say that they are not concerned with pre-theoretical intuitions about truth. There are many of these intuitions around, and it is almost certain that they cannot all consistently be accommodated. Some of our intuitions, then, will have to go, and it is the job of the philosopher to find out which.

Yet more drastically, we might even say that we are in fact dealing with a rather technical notion that need not pay any attention to intuitions at all. Then, however, it is somewhat dubious that we choose the pre-theoretically loaded term “truth” at all, instead of simply making up something new. The choice to talk of truth seems to bring with it at least *some* commitment to explicating the concept as it occurs in non-technical discourse.

A last option, of course, is to abandon truth altogether and to base the theory of meaning solely on correct assertibility. A tensed notion of correct assertibility, as I said above, does not sin against our intuitions at all. Before Fermat’s Last Theorem was proven, to assert it (rather than to conjecture it) was incorrect. Since the proof is available, it is correct to assert it.

In sum, if the constructivist can somehow get over the sheer implausibility of a tensed notion of truth, or resolve to set aside truth altogether and focus on correct assertibility, it seems that provability as given by options 1–3 has a much better chance to harmonize with Dummett’s program and also with our Kripke semantics. Now let us turn to the question which of these three options is the most promising one.

3.8.3 Varieties of Tensed Provability

Again, all that is prov*en* is surely prov*able*. To say, as option 1 had it, that *only* statements we have actually proven are provable seems a bit harsh, though. At least those statements that we know we could prove if we wanted to should also count as provable; it is hard to see any harm in that. That is, at least option 2 is unproblematic.

It also seems that the Kripke semantics would call for at least that much leeway. In fact, the value 1 corresponds to this notion of “is provable” better than “is proven” (even though “is proven” or “is verified” are the ubiquitous glosses). Every world, after all, contains all the logical consequences of what we know. It is evidently false that we *do* draw all those consequences, so if value 1 should only apply to the actually proven statements, our models would overgenerate.

On the other hand, in view of this one might want to go even further, on to option 3. That is, one might want to say that a statement is provable because a proof of it exists, and we know of a method of proving either it or its negation. A provable statement is decidable, and in fact applying the decision method, we would find a proof of it.

This way of explaining what it is to be provable fits even better with the value 1 in the Kripke semantics: It allows us to be ignorant of the truth of a complex statement, even though we know we would be able to work it out if we tried. This is necessary to actually gain any plausibility from the move from “is proven” to “is provable” as an interpretation of value 1. If we keep asking of any statement with value 1 that we know of its provability, even though we do not have to actually carry out the proof in question, then we still seem to know much more than we ever actually could at any given world.

In contrast, option 3 allows for statements that are provable, even though we do not know them to be provable. The only thing we know about them is that they are decidable. What makes them not only decidable, but also provable, is something that we do not actually know: That there exists a proof for them, given our currently available methods. But this additional knowledge would be relatively easy [22] to come by.

Not only does option 3 fit better to the Kripke semantics; if we identify provability with truth, it also satisfies our intuitions about truth (which are already strained enough by the tensedness of the account) better.

To illustrate, let us suppose that we have computed the first x digits in the decimal expansion of π. Suppose further that digit $x + 1$ is a 9. Given what we know, option 2 would say that "Digit $x + 1$ is a 9" would be unprovable and thus not true, whereas option 3 would give us that the statement is both provable and true.[23] The original intuitionists, Brouwer and his followers, might have preferred option 2. For them, the digit in question had not existed prior to our construction of it; the fact that the question whether it was a 9 or not was decidable might not have impressed them too much.

Nonetheless, I believe that option 3 is more in line with our intuitions and also with what Dummett tends to think. Take the following quote:

> [A former statement of "truth as provability"] was intended to allow as true a statement for which we have an effective decision procedure that will in fact yield a positive result, even if we do not know this. For example, a statement that a certain large prime number is prime is decidable, and may, when we apply the decision procedure, turn out to be true. (Dummett (1998), p. 123)

This clearly seems to support option 3. Even clearer is the case when we, as Dummett plans, move outside mathematics and wonder about truth as verifiability instead of provability. Suppose I say "I have fifty Euros in my wallet.", when I actually have that amount of money in my wallet.

We all know what it would take to check: just open the wallet and look. Still, if I was only estimating the amount based on the time of my last trip to the ATM and my spendings since then, and if no one else had taken a look inside my wallet recently, option 2 would say that my statement was not true. Even though everyone knows how to check, no one knows what they would find. Nonetheless, I think it is safe to

---

[22] *How* easy will depend on how much of our temporal and intellectual restrictions we are willing to idealize away.

[23] Note that both options would agree on the fact that the statement is *decidable*. The example looks deceptively similar to the one with the seventy 7s I used before to illustrate the phenomenon of undecidability. However, the important difference is that in the present case, there is a method known to us that will decide the issue, and *we know that this method will decide the issue*. In the former example, the method of simply going on to compute more and more digits was far from certain to decide the question about the string of 7s.

say that our intuition is that my statement was true, and option 3 can indeed support that verdict.

In general, the idea that an epistemically constrained notion of truth might lead to the conclusion that there are no truths that we do not know is seen as a serious drawback. Option 3 can avoid this consequence, while option 2 cannot.

Of the tensed notions, option 3 best matches common assumptions about truth. How are things with correct assertibility?

There is a clear sense in which I am right to assert a statement if I know that (and how) I could prove it, if I wanted to. So, assertibility surely at least encompasses provability in the sense of option 2. However, I think that an argument for the more inclusive option 3 can be made as well.

Let us go back to our example about the money in my wallet. If I had said that I had 50 Euros and you challenged me to prove it, then I would apply my decision method and open my wallet to see how much money was in there. Now, in terms of observable behavior, my action would not have differed, no matter whether I had actually known that that method would lead to a verification of my claim or whether I had only guessed so. I would apply the method and you would have to acknowledge the correctness of my statement.

Admittedly, this is, up to now, only backed by rather vague intuitions about what a correct assertion is. The notion of a correct assertion and its connection to verifiability and falsifiability will occupy much of our attention in later chapters.

Lastly, option 3 has an advantage that is quite independent of its fit with the notions of correct assertibility and truth and the interpretation of the Kripke models. It can suggest an answer to a challenge that Dummett put forth concerning logical consequence: If we actually have to know of every provable statement that it is provable, and if logical consequence transmits provability, then what possible use could deductive inferences have?[24] If we reason from provable premises to a provable conclusion, then options 1 and 2 will have us reason from things we knew beforehand to something we *also* knew beforehand. Why then go to the trouble of drawing the inference?

If a provable statement need not be known, on the other hand, there is a real possibility of a gain in knowledge: The relation of logical deducibility may hold between statements that we know to be true and a conclusion that only becomes known to us once we draw the inference.

All in all, option 3 strikes me as the most successful interpretation of provability, and I will understand the term in this way from now on. When I turn to notions of verifiability and falsifiability later on, they will be understood along the same lines.

[24] TOE, p. 297.

3.8.4 The Meaning of the Value 1

Let me wrap up by pointing out the precise connection of provability, as explicated by option 3, and our Kripke models. How can we understand value 1 as "is provable" in this sense?

Reading value 1 this way, we find that a statement can receive value 1 even if we do not know what we would find if we were to check. However, we need to know *how* to check. The stages are thus only in part epistemically constrained: They are so constrained in the sense that they never attach value 1 to statements that are undecidable. They are not so constrained in the sense that we have to know every provable statement to be provable.

So, a world does not in fact record our knowledge at a time, but rather the potential knowledge we could acquire without ingenuity, the things we might find out by proving or disproving every statement which is decidable. Note that the world would, as far as the Kripke semantics goes, look no different if we did. If, by moving from one stage of our investigation to the next, new statements receive value 1, then that will mean that new statements have become decidable. On the other hand, if all we have done is to prove statements that we already knew to be decidable, then no statement will change its value.

With this in mind, what are we to make of the accessibility relation? I said above that the accessible worlds mark the epistemic alternatives of a given world. In a certain sense, this is still true, but some care is needed here. Depending on how one spells out the nature of the accessibility relation, one might end up endangering the possibility of assigning value 1 to statements that will actually turn out to be false. But if this is not possible, then all negations of false statements, be they decidable or not, will receive value 1 at the actual world, contradicting the claim that only decidable statements could receive that value.

Let me unpack this. A statement will receive value 1 iff it is decidable and the decision procedure would provide a proof of it. Then, as we move to the next stage of our investigation, new statements become decidable. However, it is important that the alternatives that we can envisage include, for each undecidable statement at the actual world, worlds in which the statement becomes not only decidable, but *provable*.

So, here is a sense in which epistemic alternatives are clearly visible in the models. This is in contrast to another set of epistemic alternatives, which are *not* included in the model. These are the alternatives that we might conceive regarding decidable, but not actually decided statements. Of course, we can both imagine that there are 50 Euros in my wallet and that there are not. However, only one of these alternatives would be realized at the actual world in a model where I know of a way of finding out, but have not yet checked.

Here, then is an overview of my suggested way of interpreting provability in the Kripke semantics.

- The valuations at a world are constrained by what is decidable, not what is actually proven.

- A statement is decidable at a world if and only if either it or its negation receives the value 1 at that world.
- A statement is provable at a world if it receives value 1 at that world.
- The statements that are actually proven are a subset of the provable statements; as things stand, they are not distinguishable from those statements that are provable but not yet proven. If desired, one might add an "It is known that" operator or some other device to mark them.
- The worlds that are accessible from a world include, for every undecidable statement at the original world, worlds in which that undecidable statement receives value 1 and worlds in which it receives value 0.[25]
- Finally, logical consequence requires the following: In every model, if the premises are provable at a world, then the conclusion of a valid inference is provable as well at that world.

3.9 Chapter Summary

In this chapter, the central ideas of intuitionistic mathematics and its logic were introduced. Some of the valid and invalid inferences were discussed, and the way the meaning of the intuitionistic connectives are given in terms of their proof conditions was explained. This was the central task of the BHK interpretation, which we supplemented with the Kripke semantics to give it more traction. I spent the last sections in an attempt to give a sensible notion of constructive provability and how to interpret the Kripke semantics in terms of that notion.

So far, we have mostly been concerned with intuitionism's original subject matter, mathematics. However, as has been mentioned several times, Dummett's aim is to extrapolate the intuitionistic ideas from the mathematical realm and apply them to every kind of discourse. He thinks that this presents no substantial problem, as the argument for constructivism

> is virtually independent of any considerations relating specifically to the mathematical character of the statements under discussion. The argument involve[s] only certain considerations within the theory of meaning of a high level of generality, and could, therefore, just as well have been applied to any statements whatsoever, in whatever area of language (Dummett 1975, p. 226).

We will see how well that transition works in the second and third parts of this book. Before we get there, however, some more preparatory work is needed, which will be the business of the next chapter.

What we will learn early on in the attempt to transplant the mathematical account to the empirical realm is this: Whereas mathematics was able to do with only one central notion in its semantic theory (namely *proof*), empirical discourse might need

---

[25] In the case of mathematics, this has us assigning value 1 to statements that are mathematically impossible, which is slightly counterintuitive. Again, this sense of implausibility is assuaged by pointing out that the possibility that is reflected in the accessibility relation is an *epistemic* one.

two such notions: Verifications and falsifications. When we will turn to model logic in terms of these two notions, we will come across semantics that have truth value gaps (statements that are neither true nor false), and some even have truth value gluts (statements that are both true and false). The next chapter talks about such gaps and gluts in general, without yet tying them to verifications and falsifications in particular.

Chapter 4
Gaps, Gluts and Paraconsistency

4.1 Chapter Overview

This chapter will look at some other semantic theories and the logics they generate. Mainly, these logics come about by allowing truth value *gaps* and truth value *gluts*. If a semantic theory allows for statements that are neither true nor false, then it allows for gaps. If, on the other hand, it makes room for statements that are *both* true and false, then it allows for truth value gluts.[1]

An important theme, closely connected to truth value gluts, will be *paraconsistency*. Any[2] logic that does not allow inferences of the form $A, \neg A \vDash B$ (or, if it is syntactically defined, of the form $A, \neg A \vdash B$) is paraconsistent. This rejected rule, which we have already met under the name *Ex Contradictione Quodlibet*, is called by paraconsistent logicians, more dramatically, *Explosion*. It is valid, of course, in classical logic, but many non-classical logics are *explosive* in this sense as well. In particular, we have seen that intuitionistic logic is explosive (cf. Axiom 10 of Heyting's axiomatization).

Paraconsistency, to this day, is often confused with *dialetheism*. While the main idea behind paraconsistency is simply that a contradiction should not entail everything whatsoever, dialetheism is a much more ambitious and contentious metaphysical stance: It is the view that there actually are contradictions that are true, statements that are both true or false. I will try to clear up the relation between dialetheism and paraconsistency. Also, a lesser known alternative to dialetheism, *analetheism*, will be presented.

A further theme that will start to emerge here and continue throughout the book is the phenomenon of *duality*. Many of the concepts are mirror images of each other, but often it will matter from which direction one looks into the mirror.

---

[1] Sometimes I will talk about "gappy" and "glutty" theories. The first kind are also known as "partial" theories.

[2] One common misconception is that paraconsistent logic is one particular logical system. In fact, there are a lot of them, cf. Priest (2003).

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_4,

Finally, I will ask how these semantic theories fit into the Dummettian setup outlined in the first chapter.

Talking about gaps will give me occasion to present the distinction he draws between assertoric content and ingredient sense.

All of this material, even if this might not yet be apparent in some cases, will be important background knowledge for the remaining chapters. As most of the discussion of gaps and gluts takes place outside of expressly constructive contexts, I will mostly talk about truth and say little about assertibility in this chapter. The question how the ideas we will meet here will come to connect up with correct assertibility and verifiability later on is discussed in the last section of this chapter.

I will start the story with a simple *relevant* logic called First Degree Entailment (FDE). This logic contains all the main ideas I just mentioned, and it is easy to modify it to obtain other paradigmatic logics, such as Strong Kleene (K_3) and Priest's Logic of Paradox (LP).

Even though the logics I will eventually propose will not be relevant logics (I will explain what that means in a minute), many of the ideas that I will draw on come out of the literature on relevant logic.

4.2 Relevant Logic

What then is the concern of the relevant logician? She is unhappy about classical logic, because it allows inferences that seem to be blatant *non sequiturs*. Take the inference we just mentioned above, *Explosion.*

Isn't there something very strange about the inference "It is raining and it is not raining, therefore bananas can be used as a substitute for onions in most recipes" (an instance of *Explosion*)? The diagnosis of the relevant logician is that in an inference, the premises should have something to *do* with the conclusion and that in particular whether or not it is raining is *irrelevant* for determining the possible culinary uses of bananas. Relevant logic rejects such irrelevant inferences. So, every relevant logic is paraconsistent, because it will reject *Explosion*.

However, Explosion is not the only kind of inference that is irrelevant according to the relevant logician.[3] Consider a form of the LEM: $B \vDash A \vee \neg A$. An instance would be "The onions in this dish taste a bit mushy, therefore it either rains or it doesn't." This is obviously just as irrelevant as the first example.

It is interesting to note that both relevant logicians and intuitionistic logicians want to reject this inference, but for completely different reasons.

The intuitionist is, as we have seen, concerned about the possibility of undecidable statements.

---

[3] That is, relevant logics form a proper subset of the set of paraconsistent logics. This shows that the not uncommon perception that paraconsistency is a more radical doctrine than relevance is completely unfounded. It rests, again, on the common confusion between paraconsistency and dialetheism.

On the other hand, the relevantists usually do not care that much about decidability, nor do they necessarily take issue with the logical validity of the LEM. In fact, some relevant logicians even endorse the LEM in the form $\vDash A \vee \neg A$. What they dislike is the claim that such a tautology should follow from something completely unrelated, as in the example above, $B \vDash A \vee \neg A$.

What they want to argue is this: The meaning of the turnstile is not exhausted by "technical device that guarantees the transmission of truth from premises to conclusions." No, it is meant to be an analysis of "therefore" and similar locutions, and as such, it needs to do more than merely guarantee truth preservation.

However, if there are no premises, "therefore" makes no sense; "Therefore, it is raining or it is not raining" is only an admissible thing to say if it follows another sentence that specified some premises. If such premises are truly lacking, then some other locution has to be found to pronounce the turnstile, maybe "it is logically true" or some such.

Thus, the relevantist claims, "Blue is a color, therefore the president will be re-elected or not" is objectionable, while "It is logically true that the president will be re-elected or not" might be acceptable.

Let us see how the relevant ideas are put into practice by way of a concrete example.

4.3 First Degree Entailment

N. Belnap introduced FDE in two influential papers in the 1970s. One is entitled "A useful four-valued logic" Belnap 1977a), the other "How a computer should think" (Belnap 1977b). The two papers have been merged into Chap. 81 of Anderson et al. (1992), and this is the most accessible source.[4]

The main concerns driving this logic are first to give a basis for relevant entailment and second, as is obvious from the title of the second paper, to give computers some rules for processing information.

To see how FDE manages to get rid of irrelevant inferences, let me first remind you why classical logicians endorse *Explosion* and the LEM. Logical consequence is defined in terms of truth preservation in all models: If the premises are true, then so is the conclusion. Since contradictions have no classical models, the condition for the validity of $A \wedge \neg A \vDash B$ is vacuously satisfied. A good way to expulse this inference, then, is to provide models in which both A and $\neg A$ are satisfied. Likewise, the trick to invalidate the inference $B \vDash A \vee \neg A$ is to provide models in which neither A nor $\neg A$ is satisfied.

FDE achieves this by introducing two new truth values in addition to the classical $\mathcal{T}$ (True) and $\mathcal{F}$ (False). These values are called $\mathcal{B}$ (Both) and $\mathcal{N}$ (Neither).

[4] A scan of the chapter is available for download on Belnap's homepage, http://www.pitt.edu/belnap/papers.html.

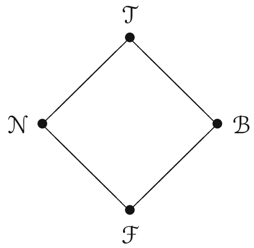
Logically complex statements get their values according to the following truth tables for negation, conjunction, and disjunction.[5]

$\neg$	
$\mathcal{T}$	$\mathcal{F}$
$\mathcal{B}$	$\mathcal{B}$
$\mathcal{N}$	$\mathcal{N}$
$\mathcal{F}$	$\mathcal{T}$

$\wedge$	$\mathcal{T}$	$\mathcal{B}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{T}$	$\mathcal{T}$	$\mathcal{B}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{B}$	$\mathcal{B}$	$\mathcal{B}$	$\mathcal{F}$	$\mathcal{F}$
$\mathcal{N}$	$\mathcal{N}$	$\mathcal{F}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$

$\vee$	$\mathcal{T}$	$\mathcal{B}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{T}$	$\mathcal{T}$	$\mathcal{T}$	$\mathcal{T}$	$\mathcal{T}$
$\mathcal{B}$	$\mathcal{T}$	$\mathcal{B}$	$\mathcal{T}$	$\mathcal{B}$
$\mathcal{N}$	$\mathcal{T}$	$\mathcal{T}$	$\mathcal{N}$	$\mathcal{N}$
$\mathcal{F}$	$\mathcal{T}$	$\mathcal{B}$	$\mathcal{N}$	$\mathcal{F}$

Algebraically, these four values can be arranged nicely in the following lattice.[6]



We will view conjunction as the *meet* (the greatest lower bound), disjunction as the *join* (the least upper bound), and negation as an operator that flips $\mathcal{T}$ and $\mathcal{F}$ but is a fixed-point operator for $\mathcal{B}$ and $\mathcal{N}$. That is to say, applying the negation operator to $\mathcal{B}$ will give $\mathcal{B}$ again and likewise for $\mathcal{N}$.

Here is what that means: Note that the four values in the diagram are connected by lines that represent an ordering relation. One value is said to be greater than the other if it is higher up on the page, and there is an ascending line from the lower value to the higher one. To compute disjunctions, one looks for the lowest value that is greater or equal than either value of the two disjuncts. For example, the value of $T \vee \mathcal{B}$ is $\mathcal{T}$ because it is greater or equal than either $\mathcal{B}$ or $\mathcal{T}$, and there is no smaller value that meets this description. How about $\mathcal{N} \vee \mathcal{B}$? There is only one value that is greater than or equal to those two values, and that is $\mathcal{T}$.

[5] There is no table for the conditional, simply because FDE does not have a conditional. This is actually the feature that gives First Degree Entailment its name: An entailment of the first degree is one in which the turnstile ($\vdash$) is the only entailment or conditional-like item. We will later see an example for how a conditional can be added to FDE.

[6] It is not essential to know what a lattice is to understand the following. Those who know what this means might benefit from knowing that the lattice is one of the de Morgan lattice variety. It also is the paradigmatic example of a *bilattice*, which in essence means that you can find two distinct lattice structures on the elements (although there are some more requirements, cf. Fitting (2002)). Additionally to the ordering that goes from bottom to top, you find the ordering that goes from left to right, that is, $\mathcal{N}$ is the lowest element and $\mathcal{B}$ the highest. The first is called the truth ordering (ascending the ordering means either gaining in truth or waning in falsity), the second the information order. Moving from left to right means to increase the amount of information available. This will become clearer in the discussion of the intuitive interpretation of the values below.

Conjunction looks for the greatest value that is less than or equal to the values of both conjuncts. For example, $\mathcal{N} \wedge \mathcal{B}$ is $\mathcal{F}$, because no other value is less than or equal to either $\mathcal{N}$ or $\mathcal{B}$. It should now be clear how the truth tables above correspond to the lattice operations.

As for negation, imagine a horizontal axis going through the values $\mathcal{N}$ and $\mathcal{B}$. Negation then is an operation that flips the value over this axis. In other words, it takes $\mathcal{T}$ to $\mathcal{F}$ and vice versa, but leaves $\mathcal{B}$ and $\mathcal{N}$ as they are, just as is recorded in the truth table.

4.3.1 Designated Values

I have told you what semantic values FDE ascribes to statements ($\mathcal{T}$, $\mathcal{N}$, $\mathcal{B}$, and $\mathcal{F}$) and how the value of a complex statement depends on the values of its parts. According to Dummett's plan, the next step in the development is to explain how the values relate to truth, that is, how we can move from the knowledge of the semantic value of a statement to the knowledge whether it is true or not. Then, we can finally find out what the consequence relation will look like, because we can define it as the relation that transmits truth from the premises to the conclusions in all models.

However, in the study of many-valued logics such as FDE, the normal procedure is often slightly different. The detour through the concept of truth is either left implicit or not intended at all. Instead, consequence is defined by singling out *designated values*, that is, values that the consequence relation is required to preserve. Whether or not these values jointly make up truth or some other desirable property is then either left open or addressed as an afterthought.

Which values did Belnap designate? Actually, he gave two equivalent choices: One may designate either $\mathcal{T}$ and $\mathcal{B}$ or one may choose $\mathcal{T}$ and $\mathcal{N}$. I will discuss the different intuitive ideas that back these alternatives below. Let us for now go with the more usual pair, $\mathcal{T}$ and $\mathcal{B}$.

Now, given these two truth values as designated, we can see how both LEM and Explosion can be avoided. The inference from C to $A \vee \neg A$ is not valid, because a counter model can be defined thus: Take C to be $\mathcal{T}$ and A to be $\mathcal{N}$. Then, the premise has received a designated value, but the conclusion has not (as will be easy to check with either the aid of the tables or the lattice diagram). On the other hand, a counter model to $A \wedge \neg A \vDash C$ can be given by assigning A the value $\mathcal{B}$ and C the value $\mathcal{F}$.

Indeed, FDE manages to avoid any inference that does not meet the *parameter sharing* requirement. This requirement is one way of making the somewhat vague notion of relevance more precise. An inference of propositional logic meets this requirement if at least one of the propositional parameters (these are the atomic statements) occurs both in the conclusion and in one of the premises. Obviously, both

LEM and Explosion fail to meet this requirement.[7] To see why all other inferences that fail to meet the requirement will have a counter model in FDE, observe that it will be possible to assign the value $\mathcal{B}$ to every atomic statement in the premises and the value $\mathcal{N}$ to every atomic statement that occurs in the conclusion. This will be an admissible valuation, because, by assumption, no atomic statement occurs both in the premises and in the conclusion, so no conflict can arise.

The next step is to note, by inspecting the truth tables, that every complex statement made up solely of atomic statements with the value $\mathcal{B}$ will receive the value $\mathcal{B}$ as well. The same goes for the value $\mathcal{N}$. This means that, no matter their logical form, the premises will receive the value $\mathcal{B}$ under our interpretation and the conclusion the value $\mathcal{N}$. But that means that we have constructed a counter model, as the valuation is one under which all premises are assigned a designated value, while the conclusion has a value that is not designated. In particular, this also implies that there are no logical truths in FDE. Any purported tautology will be seen to receive value $\mathcal{N}$ if all the atomic statements occurring in it are assigned value $\mathcal{N}$.[8] Likewise, there are no logical falsehoods in the sense of formulas which imply every other formula. It is easy to give, for any particular formula, a valuation that assigns to it the designated value $\mathcal{B}$. Again, this is achieved by setting all the atomic statements to $\mathcal{B}$. But a formula that takes a designated value in some model will not imply everything.

Finally, it should be clear that the same kinds of arguments as in the last paragraphs can be run if we go with Belnap's second suggestion regarding designated values. That is, if, instead of $\mathcal{T}$ and $\mathcal{B}$, we choose $\mathcal{T}$ and $\mathcal{N}$ as designated values. To see this, simply switch all occurrences of "$\mathcal{B}$" and "$\mathcal{N}$" in these arguments. Indeed, all inferences that are valid under one choice of designated values are valid under the other as well.

4.3.2 Thinking Computers

However, for all this to be more than a mere mathematical trick to fulfill the parameter sharing requirement, it would be nice to be given an informal interpretation of what those four values are supposed to *mean*. (Of course, the way they were named gives a pretty good hint already.)

This is the point where the second motivation for Belnap's logic comes into play. He thinks of the valuations as recording what information a computer has received about different statements. The computer is given input in the form of statements that are designated as true or false. As many people are supposed to be building up the database of the computer, it is not impossible that one person might enter a statement as true, while another enters it as false. We then end up with four possibilities for each statement:

---

[7] Inferences such as $A \wedge \neg A \vdash A \wedge B$ or $C \vdash C \wedge (A \vee \neg A)$ show that parameter sharing is necessary, but not sufficient for a relevant consequence relation. Note, however, that these inferences are invalid in FDE as well.

[8] That means that FDE will not suit those relevantists who would like to keep some tautologies, but only reject that these tautologies follow from an arbitrary premise.

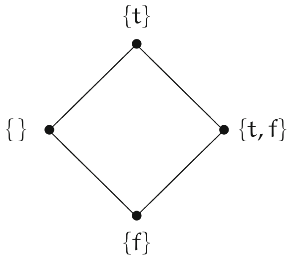
$\mathcal{N}$: The computer has received no information pertaining to the statement.
$\mathcal{F}$: The computer has received the information that the statement is false.
$\mathcal{T}$: The computer has received the information that the statement is true.
$\mathcal{B}$: The computer has received the information that the statement is true and the information that it is false.

The job of the computer now is to compute the values of complex statements and draw suitable inferences. If you prefer a less gadgety example, you might instead want to think about a criminal trial (this is not Belnap's example). Let us suppose that only testimonial evidence is available, then the transition is completely straightforward; just substitute "court" for "computer." I will come back to the example of legal trials repeatedly in the remainder of the book.

In either case, it is not clear that these four values need to have much to do with any substantial notion of *truth* at all.[9] The computer might have been fed false data, and the court might have been lied to. Belnap sometimes calls his semantic values "epistemic" values and marks the distinction between them and what he calls "ontological" truth.

Given these interpretations for the four values, how can Belnap justify the choice of designated values as being either $\mathcal{T}$ and $\mathcal{B}$ or, alternatively, $\mathcal{T}$ and $\mathcal{N}$?

To make his argument more conspicuous, it is convenient to relabel the four values in a manner suggested by J.M. Dunn. The four corners will be furnished by members of the power set of the two basic semantic values t and f. These are what Belnap calls "told-truth" values, "told-true," and "told-false." $\mathcal{N}$ corresponds to the empty set { }, $\mathcal{F}$ to {f}, $\mathcal{T}$ to {t}, and $\mathcal{B}$ to {t,f}.



The idea is not so much that there are again four semantic values that merely have different names. Rather, this way of looking at the logic suggests that there are only two basic semantic values, just like in classical logic (modulo the "told-" prefix). However, valuations are not, as in classical semantics, total functions, but rather *relations*. That is, statements will not be assigned one and only one value, but may receive none, one or two of the values on offer (told-true and told-false).

Whether the semantics is a four-valued functional one or a two-valued relational one does not make too much of a difference when it comes to formal properties. The consequence relation that will be induced will be the same one, provided the

[9] For this reason, I try to avoid calling them "truth values."

corresponding values are chosen. In a sense, only the relational semantics really has gaps and gluts, because in the functional semantics, every statement receives a value. But the values $\mathcal{N}$ and $\mathcal{B}$ are, at the end of the day, just designations of gaps and gluts as well. However, the two-valued semantics is better suited to Belnap's choice of designated values, while the four-valued version makes a different choice that I want to suggest below more comfortable to discuss.

Belnap has two distinct stories to tell about logical consequence. First and most common is the request that the consequence relation preserve *truth*, just what we have seen in the discussion of Dummett's theory. What might truth be in this setup, though? Of the two basic values, t and f, only t even comes close to any notion of truth, even if it is not "ontological" truth. As in the set formulation there are two sets that contain t, namely $\{t\}$ and $\{t, f\}$, it seems quite natural to take these as designated values ($\mathcal{T}$ and $\mathcal{B}$ in the four-valued version).

However, there is also a second property that Belnap suggests we might want to preserve, and this is something that we have not seen in Dummett (yet). This alternative is that we should like the consequence relation to transmit *non-falsity*: If the premises are not false, then neither is the conclusion. Belnap gives relatively little by way of motivation why we should call such a relation logical consequence, but we will come to discuss this topic at length in the following chapters.

Looking up the truth values that would get designated under this requirement, we find that $\{t\}$ and $\{\}$ are the only ones that do not contain value f. These values, of course, correspond to $\mathcal{T}$ and $\mathcal{N}$, and we have seen above that this choice of designated values leads to the same logic as the choice of $\mathcal{T}$ and $\mathcal{B}$.

4.4 Exactly True Logic

Very well. Either of the two features that one might wish to preserve, told-truth or non-told-falsity,[10] leads to the same logic. Is there anything more one could ask for?

That this question should be answered positively is suggested in Pietz and Rivieccio (2013).[11] Their paper argues that we should ask for a consequence relation that preserves *truth-and-non-falsity*. This property is held together with hyphens because at one point Belnap describes FDE as preserving truth *and* non-falsity,[12] which is completely correct in the following sense: If all premises are true, so is the conclusion, and if all premises are non-false, so is the conclusion.

The requirement Pietz and Rivieccio propose, however, is as follows: If all premises are true and not false, then so is the conclusion. Simply put, "told-truth"

---

[10] I will leave off the "told-" prefixes for the rest of this section to increase readability.

[11] The first author of that paper and the author of the present study are the same person, different last names notwithstanding.

[12] "Now for an account which is close to the informal considerations underlying our understanding of the four values as keeping track of markings with told True and told False: say that the inference from A to B is valid, or that A entails B, if the inference never leads us from told True to the absence of told True (preserves Truth), and also never leads us from the absence of told False to told False (preserves non-Falsity). Given our system of markings, this is hardly to ask too much." Anderson et al. (1992), p. 519.

is good, “told-false” is bad[13]; the prudent computer should choose those pieces of information that are univocally supported and infer conclusions that are similarly univocally supported. In searching material to draw inferences from, it should stay away from those inputs that it has been told are half-false (those with the value {t, f}). Of course, the only truth value that both contains t and does not contain f is {t}.

Talking “across” the two variants of semantics (the four- and the two-valued one), the difference between Belnap and the new proposal is that he wants to designate t, while Pietz and Rivieccio want to designate $\mathcal{T}$. On page 512 of Anderson et al. (1992), Belnap discusses the difference between the two values and suggests to read t as “told at least true” and $\mathcal{T}$ as “told exactly true” in circumstances where confusion between the two threatens. In view of this, Pietz and Rivieccio call the new logic “Exactly True Logic” (ETL)[14].

What happens to logical consequence if we designate only $\mathcal{T}$? For one thing, a contradiction will now never take a designated value. Therefore, the new logic validates Explosion (and thus fails to fulfill the needs of relevant logicians).

Indeed, one might well think that the new logic will coincide with a known logic. This, however, is not the case. Even though ETL validates Explosion, $A \wedge \neg A \vDash C$, just as strong Kleene does, the inference $(A \wedge \neg A) \vee (D \wedge \neg D) \vDash C$ fails. For a counterexample, take $v(A) = \mathcal{B}$, $v(D) = \mathcal{N}$, and $v(C) = \mathcal{F}$. It is easy to check that under this valuation, the premise will be assigned value $\mathcal{T}$.[15]

This is a most unusual feature. For one thing, it allows for theories that contain disjunctions, but cannot consistently be expanded by *either* disjunct, a property Pietz and Rivieccio dub *anti-primeness*.

Pietz and Rivieccio make no pretensions that these uncommon features are particularly *desirable* in a logic. However, the paper argues that the cause of the problem here is not so much the choice of designated value, but rather the logical lattice itself:

> [T]hese are quite counterintuitive features. However, when it comes to a direct comparison between FDE and the new logic, we believe that this should not weigh too heavily against the latter. This is because what we see here is merely a slight exacerbation of an unintuitive feature that has been with FDE ever since it was proposed. The lattice will give out the value $\mathcal{T}$ for a disjunction of two statements with the values $\mathcal{B}$ and $\mathcal{N}$ (...). In particular, the fact that a contradiction with the value $\mathcal{B}$ and a contradiction with the value $\mathcal{N}$ will receive value $\mathcal{T}$ when disjoined is a feature of the logical lattice, not of ETL in particular. (Pietz & Rivieccio 2013, p. 134)

[13] Bad for the proposition in question or for you, if you are reluctant to give up your belief in it. Belnap writes: “We note that in the logical lattice, each of the values None and Both is intermediate between $\mathcal{F}$ and $\mathcal{T}$, and this is as it should be, for the worst thing to be told is that something you cling to is false, simpliciter. You are better off (it is one of your hopes) either being told nothing about it or being told both that it is true and also that it is false; while of course best of all is to be told that it is true with no muddying the waters.” Anderson et al. (1992), p. 516.

[14] The logic has independently been described in Marcos (2011).

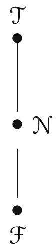
[15] This shows that the rule of proof “If $A \vDash C$ and $D \vDash C$, then $A \vee D \vDash C$” fails, which seems to stand in the way of a natural sequent calculus for this logic; the paper gives a Hilbert-style proof system instead.

In a sense, the symptoms are more visible with the new logic, but the root cause of the problem is shared by both FDE and ETL. The contrast between FDE and ETL will come to play a role in Chap. 8 when I discuss the possibility of verification–falsification gluts and the best way to handle them.

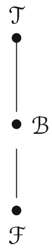
4.5 LP and K_3

Let us now move from the logics FDE and ETL with their gaps and gluts to logics that have only one of those features. The logics we will be looking at are easily obtained from FDE by tightening the requirements on the valuations. To get our gappy logic, which is known as strong Kleene logic (K_3), we simply disallow the value $\mathcal{B}$. In terms of the alternative two-valued account, we require a valuation function again, not a relation. A function is different from a relation in that it will assign at most one value to each argument. That is, no statement will be assigned more than one value, even though we are open to the possibility that it should receive none. On the other hand, our paradigmatic paraconsistent logic, G. Priest's Logic of Paradox (LP), comes about by dropping the value $\mathcal{N}$ or alternatively by requiring the valuation relation to assign at least one of t or f to each statement.[16] Algebraically put, we end up with two lattices that look quite alike and give rise to truth tables that look very similar as well:

K_3:



LP:



K_3:

$\neg$	
$\mathcal{T}$	$\mathcal{F}$
$\mathcal{N}$	$\mathcal{N}$
$\mathcal{F}$	$\mathcal{T}$

$\wedge$	$\mathcal{T}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{T}$	$\mathcal{T}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$

$\vee$	$\mathcal{T}$	$\mathcal{N}$	$\mathcal{F}$
$\mathcal{T}$	$\mathcal{T}$	$\mathcal{T}$	$\mathcal{T}$
$\mathcal{N}$	$\mathcal{T}$	$\mathcal{N}$	$\mathcal{N}$
$\mathcal{F}$	$\mathcal{T}$	$\mathcal{N}$	$\mathcal{F}$

[16] Here, there is a relevant difference between the two ways of giving the semantics. For we are dealing with a logic that, on the first interpretation, is a three-valued logic and therefore is not bivalent. However, one might argue that bivalence holds on the second interpretation, as there are only two truth values and every statement is either true or false. The only difference to classical logic is that the "either true or false" is an inclusive disjunction. Of course, one could hold that part of the idea of bivalence is that there shall be no gluts, that is, that the disjunction is an exclusive one. It is hard to guess what Dummett would have said, at least it does not clearly emerge from his extended discussion of terminology in the preface of TOE (p. xix).

On the other hand, K_3 is not bivalent, no matter how we choose to give the semantics.

Indeed, the only thing that makes a real difference for logical consequence is that in the first case, the middle value is standardly not designated, while in the second case, it is.

Unlike FDE, neither K_3 nor LP has the parameter sharing property. As is easy to see, K_3 validates Explosion, while LP validates the LEM. Indeed, LP validates *all* classical tautologies,[17] while K_3 agrees with classical logic on *all* logical falsehoods (understood as statements that imply everything).

That is to say that as the basis for a relevant logic, neither K_3 nor LP will do. There have to be other arguments for adopting such gap-only or glut-only logics. I will go through some of them quickly, focusing on those issues that will prove important to the further unfolding of the book. Note that in most of these proposals, the semantic values are not taken to be mere "told-truth" values but aspire to be genuine truth values (in one sense or other).

4.6 Uses of Gaps

Even if K_3 is not a relevant logic, there have been several applications suggested for it and similar gappy (three-valued, partial) logics.[18] Here are some of them:

- Kleene originally introduced it to deal with functions that were not everywhere defined.
- Łukasiewicz had much earlier proposed a logic with slightly different truth tables to account for future contingents.
- The semantic paradoxes and the puzzling phenomenon of linguistic vagueness have been treated (but hardly cured) with partial logics.
- Non-referring singular terms have been argued to give rise to truth value gaps.

Let us examine the last item a bit more closely, because here Dummett's attitude toward gappy theories comes out relatively clearly and because this gives me occasion to introduce his distinction between assertoric content and ingredient sense.

4.6.1 Presupposition Failure

The problem of presupposition failure is epitomized by B. Russell's classic example:

> The present King of France is bald.

[17] Including $(A \wedge \neg A) \rightarrow B$, if the arrow is interpreted as the material conditional. This makes it quite obvious that modus ponens is not a valid rule for the material conditional in LP, and usually, LP is thought of as having no conditional (not even a defined one) at all.

[18] Cf. for example Blamey (1986).

This sentence seems to say of someone that he is bald. But who is it talking about? The non-existent King of France? And which truth value should we think it has? It surely is not true, so, under the classical assumption of bivalence, it must be false. But then, shouldn't its negation be true? That is, shouldn't

The present King of France is not bald.

be true? But this seems as untrue as the first sentence!

Russell suggested that the actual logical form of such sentences is quite a bit more complicated than meets the eye, and he gave a very influential but fairly elaborate analysis that allowed the ascription of the value "False" to both of the examples without any breach of classical doctrine. The regimented but not yet formalized versions of the two statements he proposed are "There is exactly one person who is presently King of France and that person is bald" and "There is exactly one person who is presently King of France and that person is not bald."

P. Strawson, on the other hand, held that these sentences have just the logical form that you would expect [viz., Bald(King of France)] and that they are not false at all.[19] Of course, he did not suggest they are true, either; they are *neither true nor false*, that is, prime examples of truth value gaps.

Strawson argues that there is a *presupposition* that has to be met for a statement of this form to be true or false: the presupposition that the singular term in it actually has a reference. If that presupposition fails, the statement will be neither true nor false; it will either have no truth value at all or have a third truth value, neither-true-nor-false; again, the difference is relatively insubstantial. On either reading, bivalence does not hold any more.

Now, Dummett argued that this rejection of bivalence, as opposed to his own, is not a *deep*[20] one. Attributing a third truth value or a truth value gap to a statement does in and of itself not make a difference to our linguistic usage of this statement itself. What possible difference to our use of a statement would it make to call it neither-true-nor-false, instead of just plain false? Would not the speaker who asserted the statement be equally wrong in both cases?

It would of course make a difference to how we use the negation of that statement: In the case the statement that is negated is false, this negation will be true, and a speaker would be right to assert it. In the other case, where the statement is deemed neither-true-nor-false, it would be just as wrong to assert the negation of the statement.[21] A rejection of bivalence in order to "give a smooth account of the internal structure of our sentences"[22] is what he at one point thought to be a shallow one.

An interesting difference between "shallow" and "deep" (i.e., intuitionistically motivated) rejections of bivalence is that in the first case, bivalence is actually *denied*. That is, we can point to specific cases in which bivalence fails, such as statements

---

[19] Strawson (1950).

[20] TOE, p. 23.

[21] TOE, p. 12.

[22] TOE, p. xviii.

with existential presupposition failure. On the other hand, the deep reasons to reject bivalence outlined in the first chapter will not allow such counterexamples. This is because on the intuitionistic understanding of negation, being able to ascertain that a sentence is not true is tantamount to ascertaining that it is false. Thus, no statement can be known to be neither true nor false. In Dummett's terminology, his is an attack on the principle of bivalence (every statement is either true or false), while the presupposition theorists also attack tertium non datur, which says that no statement is neither true nor false.

In TRUTH, Dummett tried to defend tertium non datur by suggesting to distinguish, not between false statements and statements that are neither true nor false, but rather between different *ways* in which a statement can be false. The middle value in the truth tables for K_3 (or some other three-valued logic) would then be interpreted as a special kind of being false. Statements that are false in this particular way are such that their negations will also be false in this particular way.[23] That is, the way in which "The present King of France is bald" is false is such that "The present King of France is not bald" is false as well.

4.7 Designated Values, Assertoric Content, and Ingredient Sense

According to Dummett, then, truth and falsity *simpliciter* correspond to the class of designated and undesignated values. He asks us to appreciate the following points:

> (i) The sense of a sentence is determined wholly by knowing the case in which it has a designated value and the cases in which it has an undesignated one.
>
> (ii) Finer distinctions between different designated values or different undesignated ones, however naturally they come to us, are justified only if they are needed in order to give a truth-functional account of the formation of complex statements by means of operators,
>
> (iii) In most philosophical discussions of truth and falsity, what we really have in mind is the distinction between a designated and an undesignated value, and hence choosing the names 'truth' and 'falsity' for particular [ones amongst the] designated and undesignated values respectively will only obscure the issue. (TOE, p. 14)

Later,[24] he came to express the idea in terms of a new distinction, the distinction between the *assertoric content* and the *ingredient sense* of a statement. The first of these refers to the content of the statement on its own. It will only need to delineate those states of affairs in which an assertion of it would be correct, that is, where the asserted statement receives a designated value, no matter which.

The ingredient sense contains all that a statement can contribute to the assertoric content of complex statements containing it (see (ii) above). This might be much

[23] TOE, p. 14.

[24] For example in LBM, pp. 47–49.

more intricate than the assertoric content.[25] We need to know the ingredient sense of a statement to judge whether a complex statement of which it is a part is correctly assertible.

And here now is the connection between the designated values and assertoric content/ingredient sense:

> One way to understand the traditional semantics for many-valued logics, with its distinction between designated and undesignated values, is to take the assertoric content of a sentence to be given by the condition for it to have a designated truth-value, while the distinctions among different undesignated values, and those (if any) among different designated ones, serve to explain the ingredient senses of sentences. (Dummett 2004, p. 34)

By associating truth with the possession of a designated value and falsity with the possession of an undesignated one, he hopes to be able to acknowledge Strawson's point without having to give up his contention that the positing of gaps is not a good, deep reason to give up bivalence. What Strawson calls a gap, Dummett calls a form of falsity.

However, it is not really clear how one should incorporate the idea into the semantical account of intuitionistic logic. The Kripke semantics we have seen in the last chapter leaves no room for sentences that are unassertible because they have an existential presupposition that fails. Such statements should never receive value 1, but their negations should never receive value 1 either. However, we know that the negation will receive value 1 immediately as soon as we realize that the negated statement will never receive value 1.[26] It might be possible to augment the semantical theory to accommodate the idea, but Dummett does not develop a concrete proposal.

But why does he think that such a strategy to explain away truth value gaps is necessary anyway? Why not reject bivalence for whatever kind of reason there might be, whether deep (his) or shallow (Strawson's)? K. Green gives an interesting interpretation here (Green 2005). According to her, Dummett sees his rejection as deep because it has a deep metaphysical consequence: the rejection of realism.

The truth value gaps, on the other hand, might not lead to such a deep result. It seems to be a completely tenable option to hold that truth value gaps exist but have nothing at all to do with our cognitive abilities. However, Green argues, Dummett wanted the deep implications of the failure of bivalence, and giving up bivalence for the wrong (shallow) reasons would threaten those implications. Therefore, the shallow rejection of bivalence could not be allowed to stand.

---

[25] While I do not know of an a priori argument why assertoric content and ingredient sense might not turn out to be completely distinct, I would guess that in most theories, we will find the assertoric content somehow subsumed under the ingredient sense. The ingredient sense will have to answer; for example, how conjunctions are decided to be assertible, and it seems hard to answer that if we do not know the conditions under which the conjuncts were assertible on their own.

[26] However, we can see a different example in the Kripke semantics in which the assertoric content and the ingredient sense come apart: Assume that a statement receives value 0 at a world. Then, it is already settled that it is, at that world, not assertible. However, more is needed by way of information to decide whether the negation of that statement is assertible, namely the future development of our investigation.

Green goes on in her essay to observe how Dummett grew more and more lenient toward truth value gaps over time. This is because he came to see *any* kind of rejection of bivalence as a form of anti-realism (cf. LBM, p. 325). Thus, the connection between bivalence and realism is, according to Dummett's later view, not threatened by truth value gaps, and thus, there is no real dialectical need for him to oppose them.

4.8 Motivations for Paraconsistency

Let us now come to truth value gluts and paraconsistency. Just as there is a great diversity of incentives for gaps, there are many motivations for paraconsistency apart from the considerations involving relevance we saw earlier. Again, a paraconsistent logic is one in which the inference, known as *ex contradictione quodlibet* or Explosion, from a contradiction to an arbitrary statement is rejected. Indeed, we have already seen such a rejection in the last chapter, when minimal logic was mentioned (4.3.6.1).

However, there is some reluctance to classify minimal logic as a genuine paraconsistent logic. While not every statement can be derived from a contradiction, a contradiction will entail every *negated* statement. It is easy to see why this is so, if one remembers that $\sim A$ is actually short for $A \supset \bot$. A contradiction will then entail $\bot$ simply by modus ponens, and because minimal logic, just as intuitionistic and classical logic, supports the inference from C to $B \supset C$, we can draw an inference from $\bot$ to $B \supset \bot$ (i.e., $\sim B$) for any B.

For most mainstream paraconsistent purposes, this is too much; little is gained if not all statement, but all negated statements are derivable from a contradiction.

Here is a partial list of reasons for turning to a (truly) paraconsistent system[27]:

- The recognition that there are interesting but inconsistent scientific theories and the perceived need to treat them without inferring everything whatsoever in them.
- As a special case, there is the research conducted in inconsistent mathematics (Mortensen 1995).
- Paraconsistent logics have been proposed to reason about inconsistent fictions.
- There are inconsistencies in most bodies of law that need to be dealt with.
- Finally, the semantic paradoxes like the Liar and others have been a main concern of paraconsistent logic.

In the majority of cases, paraconsistency is achieved by allowing gluts in the semantics. However, the connection is not a necessary one. Below, I will introduce a view called Analetheism according to which it is actually *gaps* that induce paraconsistency.

As I said before, paraconsistency is not to be confused with *dialetheism*, the view that some contradictions are actually true. The next section gives a very quick introduction to the fascinating world of dialetheism.

[27] See, for example, Berto (2007) for more.

4.9 Dialetheism

The main reason to be a dialetheist has always been the deep puzzle posed by semantic paradoxes like the famous Liar sentence:

This sentence is false.

These paradoxes have been around for the longest time without any plausible consistent solution in sight. This is not due to a lack of trying; some of the smartest philosophers have tried hard to explain away the obvious contradictions such statements give rise to, to little effect.

The dialetheist[28] argues that this is because there is nothing to explain away: These statements are true and false, just like they appear to be.

This view has some important advantages. Here is one of them: I listed the semantic paradoxes under the motivations for partial logics as well. Here, the view would be that the Liar sentence and its kin are neither true nor false. But this view, unlike dialetheism, is open to a relatively straightforward counterargument. It might do away with the Liar, but what about the following "Revenge Liar":

This sentence is either false or it has no truth value.

It is easy to see that if one holds that this sentence has no truth value, it will be true, contradicting the view on display. Dialetheism is quite immune to such attacks.

On the other hand, dialetheism has a big disadvantage as well: It seems utterly unbelievable. It contradicts what many have seen as the most basic insight of all, the law of non-contradiction[29]:

LAW OF NON- CONTRADICTION: No statement can be true if its negation is, and no statement is both true and false.

As we see, the law of non-contradiction incorporates two very closely related ideas. Often you will find the law of non-contradiction defined as only one of these ideas. Many of these authors will understand the other principle to be entailed by the one they make explicit. I will use the above definition, and when there is occasion to discuss the two principles separately, I will adopt the following terminology:

NO GLUTS: No statement is both true and false.
and
NEGATION INCOMPATIBILITY: No statement can be true if its negation is.

The dialetheist takes both of these aspects of the law of non-contradiction to be unfounded prejudices and tries to give cogent counterarguments. It is fair to say that it took the philosophical world some time to take dialetheism seriously, but by now, it has become a major position in the discussion of the paradoxes.

---

[28] The most important exposition and defense is Priest (2006a).

[29] Unfortunately, received terminology works a bit against clarity in this case: The law of non-contradiction is a semantic principle, like bivalence, and not a logical principle like the Law of Excluded Middle. As said above, I write semantic principles in lowercase letters in the hope to avert confusion.

As far as the requirements on logic that dialetheism entails are concerned, again it is clear that paraconsistency is indispensable. Else the dialetheists would have to infer everything from the Liar sentence, which they take at face value. LP was the first suggestion that Priest made in the 1970s, and even though it went through some modifications,[30] it is still the basis of the logic he advocates today.

4.10 Expressing the Law of Non-Contradiction

A striking fact that follows from LP's validating all classical truths is that it also validates $\neg(A \wedge \neg A)$. This must come as a surprise to anyone who is told that the law of non-contradiction is supposed to fail in LP. What else but an expression of that law could the validity of $\neg(A \wedge \neg A)$ be?

LP is not the only paraconsistent logic that sports this remarkable feature. In fact, I would guess that *most* paraconsistent systems share it.[31] As long as no connection between paraconsistency and the law of non-contradiction is claimed, there is not even any particular tension here. We have no reason to view a paraconsistent logic as flawed because of its validating $\neg(A \wedge \neg A)$, because we have defined paraconsistency not in terms of this schema, but rather in terms of the failure of Explosion, $A, \neg A \vdash B$. If this specific feature fails to capture the rejection of the law of non-contradiction, then so be it.

That being said, most people in the debate nowadays *do* seem to think that it is by being paraconsistent rather than by invalidating $\neg(A \wedge \neg A)$ that a logic flaunts the law of non-contradiction, or rather that by sanctioning Explosion rather than by validating $\neg(A \wedge \neg A)$ that a logic makes its allegiance to the law of non-contradiction known. This view can be witnessed in many chapters of a relatively recent volume on the law of non-contradiction (Priest et al. 2004) (cf. especially Brady's, Restall's, and Grim's contributions).

This is a very modern view of the matter. Not too long ago, the role of logic was seen as delineating a specific set of formulas, the tautologies. On this view, there could be no doubt that the question whether a logic satisfies the law of non-contradiction is the question whether $\neg(A \wedge \neg A)$ is derivable; which other tautology should be better suited to say that there are no contradictions?

However, as more and more non-classical logics came into view, the idea of logic as only concerned with tautologies had to give way. As we noted, LP and classical logic are indistinguishable if we only look at their tautologies. But of course, they are different logics. Similarly, both strong Kleene and FDE have the same set of tautologies, namely the empty set. This obviously does not make them the same logic.

---

[30] Cf. Chap. 16 of Priest (2006a).

[31] This is impressionistic. Counting is, as so often, difficult, as there is an infinite number of different paraconsistent systems.

Logic then has to be about something more than just tautologies if we want to be able to differentiate between the above logics. A more comprehensive view is that logic is about consequence relations. The job of a logic is to separate the valid inferences from the invalid ones. This view allows us to distinguish between classical logic and LP, and similarly between strong Kleene and FDE, by noting again that (for example) $A \wedge \neg A \vdash B$ is a valid inference of classical logic and strong Kleene, while it is not in LP and FDE.

While this view of logic now makes it *possible* to say that a logic like LP, while it may allow the derivation of $\neg(A \wedge \neg A)$, is nonetheless commendable to someone who rejects the law of non-contradiction because it is paraconsistent, it is far from clear that one *should* say that. The argument normally given at this point is that if you accept contradictions, then the further contradiction between the contradiction you accept and the negation of it that you accept on account of its validity does nothing to weaken your position. That is, if you accept A and $\neg A$, then you accept a contradiction, and the further news that this compels you to accept another pair of contradictory statements, $A \wedge \neg A$ and $\neg(A \wedge \neg A)$, should not concern you. Of course, it is not just two instead of one contradiction you will have to accept, because from the second contradiction, a third is easily generated and so on. However, even in view of this staggering number of contradictions that you are committing yourself to, you still can go on reasoning as long as your logic is paraconsistent. So, the law of non-contradiction can not be captured by the validity of $\neg(A \wedge \neg A)$, or else it should prevent such ongoing use of the logic when it is applied to inconsistent premises.

This might be an argument against seeing the law of non-contradiction embodied in the validity of $\neg(A \wedge \neg A)$, but it gives no grounds yet why a friend of contradictions should want these formulas valid. Indeed, I do not think there is an argument here over and above the fact that these validities come bundled up with other logical commitments. For example, they follow from the following principles that Priest wants to endorse: the validity of Excluded Middle,[32] Double Negation Elimination, and the de Morgan laws [from $\vdash A \vee \neg A$ infer by de Morgan $\vdash \neg(\neg A \wedge \neg\neg A)$ and by Double Negation Elimination and permutation $\vdash \neg(A \wedge \neg A)$].

That is to say, surely the dialetheist should not object to, say, FDE because it does not validate $\neg(A \wedge \neg A)$, whereas he may or may not want to object to it on the grounds that it does not validate Excluded Middle (for reasons that are probably unrelated to her dialetheism). Indeed, there still seems much to be said for trying to keep the number of contradictions down and not have a single contradiction mushroom up to an infinity of contradictions in no time. But granted, a dialetheist can not be forced to renounce his dialetheias by force of the validity of $\neg(A \wedge \neg A)$ alone.

Now, what about the relation between Explosion and the law of non-contradiction? If a logic has any means at all to fend off contradictions, then it is by blowing up into the trivial consequence relation in an heroic act of suicide bombing. This is what

[32] He argues against truth value gaps and intuitionism alike in Chap. 4 of Priest (2006a), though in the auto-commentary to that chapter that is supplied in the second edition of the book, he takes a slightly more lenient approach.

one gets for bringing in contradictions into such a system, so one should better not do it.[33]

If this is the only way how the law of non-contradiction can be enforced by a logic, then giving up the law will mean giving up this defense mechanism. It need not be the other way around, though. If you have arrived at a preference for paraconsistent logics by way of concerns about relevance, for example, you have no apparent need to allow for the existence of true contradictions.

This last section is not at all an exhaustive account of the state of the discussion about the very thorny issue of the correct definition of the law of non-contradiction. It is only meant to flag the issue and warn against the very plausible, but false assumption that a paraconsistent logic could not count $\neg(A \wedge \neg A)$ among its theorems, because most paraconsistent logics that will come up in the remainder of this book will do just that.

4.11 The Law of Non-Contradiction, Bivalence, and Duality

Before leaving the subject, note the similarity between the dialetheic rejection of the law of non-contradiction and the intuitionistic rejection of bivalence and how they influence the logical principles. The law of non-contradiction and bivalence are *dual* principles: One forbids gluts, the other gaps. To those, semantic principles correspond Explosion and the LEM, which are in an important sense dual as well, although that sense is not perfectly obvious.

Intuitively, to dualize something is to "flip it over" in some way, such as when I dualize my face by looking into a mirror, or by taking a picture of it and then inverting the color spectrum in Photoshop, or maybe by pressing my head into a bowl of plaster. All of these actions leave me with some sort of dual of my face, but of course these duals look quite different.

Likewise, in logic and mathematics, the meaning of "dual" is quite context sensitive, and often one does not know what an author who uses the term means by it until one sees some examples. It can refer to the switching of polarly opposite connectives, such as conjunctions and disjunctions, necessity and possibility operators, or universal and existential quantifiers. It might involve the deletion or the insertion of negations. On the algebraic or semantical side, it can refer to the inversion on some algebraic order, the substitution of open sets for closed ones, a switch in designated values such as from true to non-false ones, or the switch from a underdetermined valuation function (one with gaps) to an overdetermined valuation relation (one with gluts). The correspondence of these switches in the semantics to the valid inference patterns is most often far from obvious.

---

[33] If even the prospect of being committed to trivialism, the view that everything is true, cannot scare you off, then even Explosion cannot compel you to keep your reasoning contradiction free. Priest has tried to argue against an imaginary trivialist in Priest (2006b), and this turns out not to be an easy task at all.

P. Halmos and S. Givant describe the potential confusion the mention of duality can cause in logic (they are writing about classical propositional logic in its algebraic guise as a Boolean algebra):

> If an experienced Boolean algebraist is asked for the dual of a Boolean polynomial, such as say $p \vee q$, the answer might be $p \wedge q$ one day and $\neg p \vee \neg q$ another day; the answer $\neg p \wedge \neg q$ is less likely but not impossible. (Halmos & Givant 1998, p. 47)

If we widen our scope and consider single premise entailments of the form $A \vDash B$, there are even more possibilities of duality, as there is the further option of a switch between the left- and right-hand side of the turnstile. The inferences we are concerned with, $B \vDash A \vee \neg A$ and $A \wedge \neg A \vDash B$, are then dual in the sense that premise and conclusion are switched, as well as conjunctions and disjunctions.

So, we observe that the rejection of the semantically dual principles of bivalence and the law of non-contradiction results in the rejection of the two inferences LEM and Explosion, which are dual in the sense just mentioned. However, had one been asked to guess the dual of $\vDash A \vee \neg A$, another plausible answer would surely also have been $\vDash \neg(A \wedge \neg A)$, which we have just seen to be valid in LP.

Likewise, a dual of $A \wedge \neg A \vDash B$ could also have been $\neg(A \vee \neg A) \vDash B$. This pattern is valid in intuitionistic logic, so here is yet another aspect in which there is a duality between LP and intuitionistic logic: The startling fact that giving up the law of non-contradiction does not entail giving up $B \vDash \neg(A \wedge \neg A)$ is mirrored by the equally surprising fact that giving up bivalence does not force us to give up $\neg(A \vee \neg A) \vDash B$.

There is a point where this kind of duality breaks down, though: Intuitionistic logic validates $\vDash \neg\neg(A \vee \neg A)$, while $\neg\neg(A \wedge \neg A) \vDash C$ is not a valid inference of LP. Not surprisingly, the duality of LP and K_3 is stronger than the duality of LP and intuitionistic logic. We will get to know a logic that has a better claim to being "the" dual of intuitionistic logic soon. It is named, aptly enough, *dual intuitionistic logic*.

4.12 Analetheism

A recent discussion that also exemplifies a form of duality is that between dialetheism and *analetheism*, a new philosophical position canvassed by J.C. Beall and D. Ripley.[34] Dialetheism, as we have seen, encompasses the idea that there are truth value gluts, that is, sentences that are both true and false, and furthermore stipulates that such sentences are assertible, since all that is true is assertible. Analetheism takes another route: In the authors' own words: "Analetheism, for us, is the thesis that some sentences lack truth-value, coupled with the willingness to assert such sentences."[35] Analetheism, thus, has assertible gaps as opposed to the assertible gluts of dialetheism. A different way of phrasing the credo of analetheism is thus: "Assert

[34] Beall and Ripley (2004).

[35] Beall and Ripley (2004), p. 30.

only that which is not false," rather than the dialetheist's "Assert only that which is true." Thus, one could argue that analetheism was foreshadowed by Belnap's idea of logical consequence as transmission of non-falsity.

As to which sentences they have in mind and how to treat them logically, Beall and Ripley follow the dialetheist's arguments closely. They principally want to address the semantic paradoxes, and they want to suggest LP as the appropriate logic, at least in terms of the consequence relation. However, they use the truth tables of K_3, that is, they regard the middle value as a gap rather than a glut. But in contrast to K_3, they take this gappy middle value as designated. The resulting logic, of course, coincides with LP. Analetheists and dialetheist are in complete agreement which statements are assertible and which are not.

Thus, analetheists and dialetheists in particular agree that there are designatedly valued (that is, assertible) statements of the form $A \wedge \neg A$, but they disagree both about the question whether these statements are true and about the question whether they are false. Thus, as Beall and Ripley observe, "each position runs counter to one traditional dogma while accepting another,"[36] the two dogmas being the unassertibility of non-truths and the untruth of contradictions. A decision between dialetheism and analetheism, to them, seems to have to be based on a decision on which of these two dogmas should be retained and which one given up, as the rest of the two theories are so similar in motivation, virtues, and vices.[37]

As they make clear, given this close proximity, they are not able to come to a conclusion which theory should be preferred and thus are not in the business of advocating analetheism over dialetheism. Rather, they point out the apparent stalemate and challenge the dialetheist to explain why she is not an analetheist. The choice seems quite arbitrary indeed, but in later parts, I will present a paraconsistent view that, if anything, is a form of constructive analetheism.

4.13 Chapter Summary

In this chapter, I have introduced the ideas of truth value gaps and truth value gluts. The first occurs if some statements fail to receive a truth value, the second if some statements receive more than one value. I gave example logics exhibiting these phenomena, the gappy K_3, the glutty LP, and the logic FDE, which has both gaps and gluts. Moreover, I gave some idea of the philosophical motivations that have been offered for these logics, and gaps and gluts more generally.

Furthermore, we have seen a number of dualities in this chapter: the duality between gaps and gluts, the duality between truth preservation and non-falsity preservation, and the duality between the LEM and Explosion. I also stressed that it is not

[36] Beall and Ripley (2004), p. 34.

[37] There is, quite obviously, also a view on which there are gluts but no gaps and according to which only non-falsities should be asserted. It should not come as a surprise that this view (which I do not think has a dedicated name) would give rise to K_3's consequence relation.

always clear what the dual of a concept, a formula, or a position is supposed to be. Is *the* dual of LP the logic K_3? Or is it the logic that uses the truth tables of LP but transmits non-falsity? Or is it the logic the analetheists prefer, which uses the truth tables of K_3 and transmits non-falsity?

This shows the difficulty of the concept of duality, a difficulty often played over in formal texts, where the reader often gets the mistaken impression that "dualization" is a clearly defined term and that its definition is common knowledge. This is not the case though, and if the definition is not given explicitly, one has to be careful to pick up the intended meaning. Duality will feature heavily in what is to come, and I shall try to be clear about what I mean when I use the word.

Having given this promise let me end this chapter by asking the following: How do the various semantic theories with their gaps and gluts fit into the role laid out for such theories by Dummett?

Again, Dummett's idea was that a semantic theory should spell out how complex statements receive their semantic values, given the values of the constituent statements. Once that was accomplished, it should show how the semantic value of a statement determines its being assertible/true or not. Given that, logical consequence can be defined in terms of preservation of correct assertibility or truth.

There was one idea in this chapter that seems not to fit particularly well into Dummett's scheme. That idea was that logic need not be defined as truth preservation, but that it might rather be defined as non-falsity preservation. This came up in Belnap's discussion and is the leading idea of analetheism. I pointed out that there is yet a third alternative here that suggests itself naturally, namely truth-and-non-falsity preservation, and showed that at least in the case of Belnap's lattice, this choice is different from the other two.[38]

The incompatibility with Dummett's program of these ideas is not too grave, though, and we will soon see that Dummett at times suggests something like non-falsity preservation as the base of logical consequence as well.

There is another way of looking at the differences here, and that is to say that the question is not how consequence is defined, but how semantic values are wired to truth. This might be less plausible in the case of the interpretations we have seen the dialetheists and analetheists give, who clearly single out some of the values as true and others as untrue. On the other hand, it might be a sensible way to view the differences between FDE and ETL. The way to report this difference would then be that FDE takes the value t to correspond to truth, while the new proposal would take truth to coincide with value T.

Be that as it may, one may well wonder what the semantics of FDE are supposed to have to do with Dummett's project at all. The semantic values are given in terms of told-truth and told-falsity. Even if there are different ways of matching these up with a concept of truth simpliciter that can then be used to define logical consequence, this concept will surely end up being one variety or other of uncertain, hearsay "truth."

---

[38] If the underlying lattice is either of the three-valued ones, the consequence relation defined in terms of truth-and-non-falsity preservation will coincide with K_3.

Certainly, this would qualify as a notion of truth that is very "anti-realistic," in the sense that it is far from the realistic notion of truth.[39] But Dummett, it would seem, had something a little more weighty in mind than just pieces of unchecked information. A proof of p, for example, certainly seems to imply something more solid than just a piece of information claiming that p. Even if we are not after truth at all, but content ourselves with correct assertibility, something more conclusive seems called for. Simply being told that p by someone seems not enough to make a correct assertion that p.

However, as I mentioned several times before, the projected transformation of the intuitionistic account of mathematical statements to an account of empirical statements will involve an important exchange of central concepts: Whereas mathematics can talk of proofs, in the empirical realm, there are *verifications* and *falsifications* to build on. The conceptual leap from *told-true/false* to *is verified/falsified* is not too great, but it will require some discussion whether we want to allow gaps and/or gluts between the latter pair.

The move from proofs to verifications and falsifications is the concern of the next chapter.

---

[39] Indeed, Wansing (Wansing 2012) offers a notion of "non-inferentialist, anti-realistic truth" based on told-truth values.

Part II
Falsifications

Introduction to Part Two

Part One supplied the necessary background information on Dummett's general program, the intuitionistic tradition he took inspiration from, and an idea of the principles of gappy and glutty logics. Building on that, I will now proceed to the main part of this work.

The main question that was left unresolved up to now is how the move from mathematical constructivism to empirical constructivism is going to be managed exactly. Of course, it is clear already that the first step is to start to talk about verifications instead of proofs. However, we will soon see that verifications alone might not be enough. More precisely, the twin notion of a *falsification* will have to be introduced.

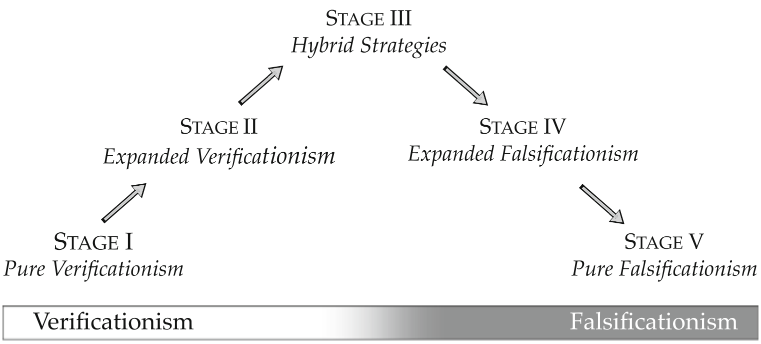
Now, we will find that there are many ways in which verifications and falsifications can be combined in a constructive theory of meaning. Which of these different options we choose will have, so I shall argue, important consequences for the logic we will want to adopt. It is the main task of Part Three of this book to investigate which constructive logics are the most suitable candidates for each variant of mixing verifications and falsifications.

This middle part is mostly here to survey the options that are available and give some of the reasons why we should consider them in the first place.

Some of these options were contemplated by Dummett himself, others were not. In the first of the two chapters that make up Part Two, I discern five main stages of falsificationistic involvement, one of which (Stage III) will fracture into yet more options before long. It is the main aim of the next chapter to clearly separate these stages.

I have organized these five stages in a peculiar pyramid structure. This pyramid has two axes that I will utilize to illustrate two different features of the theories.

First, the horizontal axis records the relative importance that verifications and falsifications play in the theory. The further to the left a position is, the heavier is the lifting that verifications have to do. The further to the right a position is, the more it relies on falsifications.



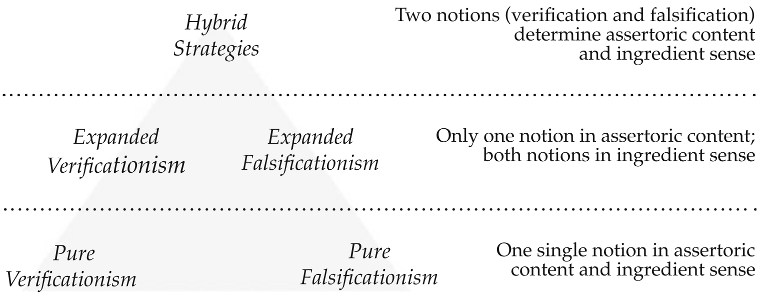
Indeed, if a position is on the left slope, I will call it a form of *verificationism*, while if it is on the right, it will be a form of *falsificationism*. Whether Stage III is a form of one or the other is not quite so easily answered, but we'll have settled this issue by the end of the next chapter.

Verificationism and falsificationism, as I will be using the terms, differ primarily in their assessment of what it is to make a correct assertion. While the idea of verificationism, as we have seen in the earlier chapters, is that an assertion is correct iff it is verifiable, the norm that falsificationism employs is this: An assertion is correct iff it is unfalsifiable. This seems a rather unnatural option, but Dummett at times argues for its superiority over the verificationistic account. I will investigate how plausible this basic idea that underlies the right-hand side of the pyramid really is. This will be the business of the third chapter of this part (Chap. 6), bearing as it does the tell-tale title "Falsificationism."

Going back to the pyramid, the vertical axis represents the amount of mingling verifications and falsifications will get to do. The lower a position is situated in the graph, the more homogenous it is. The higher up, the more interaction between the two notions there will be.

At this point, I would like to remind you of the distinction Dummett drew between *assertoric content* and *ingredient sense*. Recall that the assertoric content is what determines whether an assertion was made correctly or not. The ingredient sense, on the other hand, needs to be known in order to predict the effect the statement will have on the assertoric content of a more complex statement containing it.

With that in mind, let us look at the diagram again. In the theories at the bottom level (I and V), only one notion will be employed. These are therefore called the *pure* theories, viz., pure verificationism and pure falsificationism.



On the next tier, we find what I call the *expanded* theories: Expanded verificationism (II) and expanded falsificationism (IV). In these we will see a combination of verifications and falsifications in the ingredient sense. However, the assertoric sense is still exclusively determined by verifications (Stage II) or falsifications (Stage IV).

Finally, Stage III theories allow both verifications and falsifications in both the ingredient sense and the assertoric content. This stage will allow for more than one variation, all subsumed under the label *hybrid strategies*.

So much for the quick overview. The next chapter will go through these options much more carefully.

Chapter 5
From Proofs to Verifications, and on to Falsifications

5.1 Chapter Overview

In philosophical circles, Dummett may well be seen as the most important campaigner for intuitionistic logic in the second half of the last century. This is because, as we already know, he had a novel and most ambitious vision: He would free intuitionistic logic from the narrow confines of mathematical discourse and show its applicability to *all* of our discourse and reasoning.

The key move in this project is the replacement of the concept of *proof* in the intuitionistic account with the concept of *verification*. As we have seen in Chap. 2, a central task (maybe even *the* central task) of a theory of meaning is to account for correct assertibility. Where the mathematical intuitionist was only correct in asserting something that was constructively provable, the verificationist will make a correct assertion if, and only if, what he says is verifiable. I will call any theory of meaning that is based on this central identification a form of *verificationism*.

VERIFICATIONISM An assertion is correct iff it is verifiable.

A theory that has only the concept of verification as its central concept will be called a form of *pure verificationism*.

However, we will see how Dummett runs into a problem when he tries to translate the clause for intuitionistic negation into purely verificationistic terms. The remedy that he himself proposed was to add *falsifications* to the account.[1] That is, instead of one central notion in mathematics (proof), here there will be two important notions that have to be incorporated into a theory of meaning, viz., verification and falsification. Their relative weight will have to be assessed, and I will follow Dummett in doing just that and extend his analysis.

At first, Dummett allows for falsifications to play a role only in determining the meaning of negations, that is, as part of the ingredient sense. Although one needs to

---

[1] Although Dummett often talks about meaning theories based on verification and falsification conditions in later works, the place where he extensively discusses falsifications is his *What is a Theory of Meaning (II)*.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_5,

know both verification conditions and falsification conditions of a statement to be able to use it properly as a constituent of complex statements, the knowledge of the verification condition alone would suffice to use it on its own. The assertoric content is made up solely by its verification conditions. This is what I will call an *expanded verificationism*.

He then considers the possibility of giving the two concepts equal weight in the following way: An assertion could be held to be correct iff verifiable and incorrect iff falsifiable. This possibility, however, is soon dismissed, and I will spend some time in this and subsequent chapters to investigate whether this quick dismissal is justified. I will find that there are more options to consider here than Dummett allowed, and I will subsume them under the label *hybrid strategies*.

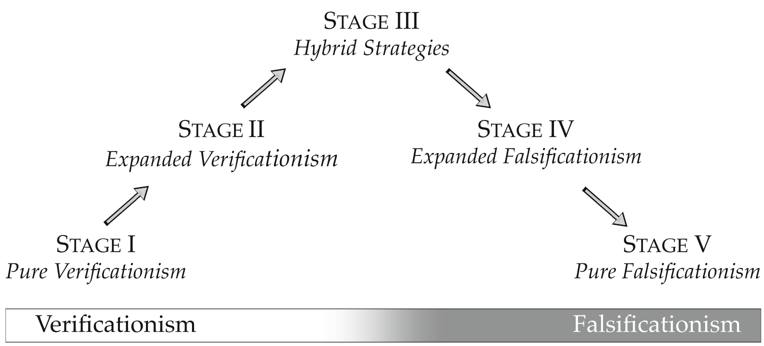
Lastly, Dummett suggests that we might be able to cope with only *one* of the two central notions after all. The thought is that it might be possible to build up a theory of meaning without any recourse to verifications at all, utilizing only falsifications. This seems to fit well to an important idea of his that we have not yet encountered: Between the notions of correct and incorrect assertions, it is actually the latter notion that is primary. Grasping the meaning of a statement is to know the conditions under which an assertion of it would be incorrect. Because he sees a close tie between finding an assertion incorrect and falsifying the asserted statement, he is driven to explore a purely falsificationistic theory of meaning. In such a falsificationism, an assertion is considered correct iff it is unfalsifiable.

FALSIFICATIONISM An assertion is correct iff it is unfalsifiable.

As his specific proposal is running exclusively on falsifications, without any need of verifications, I will call it *pure falsificationism*.

I will add one further stage that Dummett did not take into account: As it is not clear that a purely falsificationistic theory will fare much better than a purely verificationistic one when it comes to explaining the meaning of negated statements, I propose to investigate an *expanded falsificationism*. That is, of course, a theory that takes falsifications as the main notion in the assertoric content but makes use of verification conditions in the ingredient sense.

I thus discern five subsequent stages, with falsifications assuming an ever more prominent role down the line:



Stage I: *Pure verificationism*. This is the straightforward adaptation of the intuitionistic program to the empirical realm, with verifications assuming the role of proofs and no falsifications in sight.

Stage II: *Expanded verificationism*. Those are the theories of meaning that remain verificationistic in spirit, but employ falsifications in the ingredient sense, mainly to account for negations.

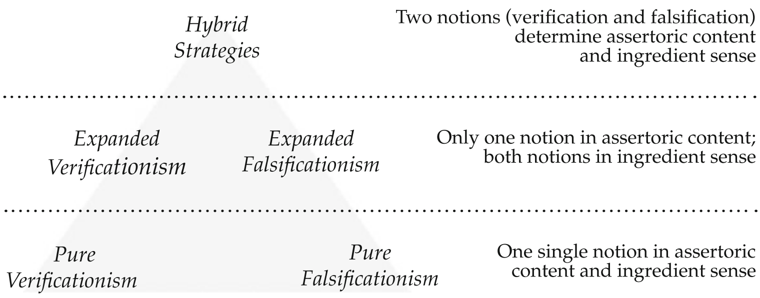
Stage III: *Hybrid strategies*. Theories in which verifications and falsifications are equally important to fix the both assertoric content and ingredient sense of statements. We will find three different approaches here, which I call the *CV&IF* strategy, the *discourse separation* strategy and the *burden of proof distribution* strategy.

Stage IV: *Expanded falsificationism*. These are theories that take Dummett's idea seriously that the incorrectness of assertions is more important than their correctness and that such incorrectness is to be tied to falsifications. However, at this stage, verifications are still around in the ingredient sense. They help to account for complex statements, just like falsifications did at Stage II.

Stage V: *Pure falsificationism*. Theories at this last stage are constructed according to Dummett's idea of how to dispense with all verifications. We find only falsifications, both in the ingredient sense and in the assertoric content.

The array above is not a flat one. The shape is, on the one hand, meant to emphasize that the two sides of the pyramid answer to two different accounts of assertibility. On the left, the verificationistic slope, a statement is assertible iff it is verifiable. The right side is the falsificationistic one, and an assertion is correct iff it is unfalsifiable.

Secondly, it will, unsurprisingly, turn out that there is a close similarity between Stages I and V on the one hand and Stages II and IV on the other. The pyramidical arrangement makes this symmetry even more obvious than the names of these stages themselves would have done.



The strategies that are found on the same level share not only the number of central concepts in the ingredient sense and the assertoric content. Also, the logics they support will turn out to be very closely related. The phenomenon of *duality* I talked about in the last chapter will make a frequent reappearance here.

Again, this chapter merely introduces the different stages; it will be the task of the chapters in the third part to shed some light on the question which logical systems are most likely to be motivated by theories of meaning at the different stages. Before that, the next chapter will investigate the unfamiliar take on assertibility that is driving the falsificationistic project.

As for the layout of the chapter you are reading now, I will mostly follow the rather obvious straight path marked by the arrows in the first diagram, which is also roughly the path Dummett took in his WTM. In Part Three, it will prove more useful to work through each horizontal layer at a time, from the bottom up.

Here, I will thus start with a short account of pure verificationism[2] and the problem that made the introduction of falsifications necessary in the first place: the meaning of negated statements.

5.2 Stage I: Pure Verificationism

To repeat, the first step in transmigrating the intuitionistic account of meaning to empirical areas of discourse is this: The mathematician's quest for proofs is replaced by a quest for *verifications*. Dummett writes:

> [The intuitionistic] theory of meaning generalizes readily to the non-mathematical case. Proof is the sole means which exists in mathematics for establishing a statement as true: the required general notion is, therefore, that of verification. On this account, an understanding of a statement consists in a capacity to recognize whatever is counted as verifying it, i.e. as conclusively establishing it as true. It is not necessary that we should have any means of deciding the truth or falsity of the statement, only that we be capable of recognizing when its truth has been established. The advantage of this conception is that the condition for a statement's being verified, unlike the condition for its truth under the assumption of bivalence, is one which we must be credited with the capacity for effectively recognizing when it obtains; hence there is no difficulty in stating what an implicit knowledge of such a condition consists in—once again, it is directly displayed by our linguistic practice. (WTM, p. 70)

He is putting forth, in short, a verificationistic theory of meaning. The content of a statement, so the verificationistic theory of meaning, is given by what would verify it. A statement is correctly assertible iff it is verifiable.

In talking about *verifiability*, we encounter a similar broad range of possible interpretations as in the case of *provability*. Although it seems that Dummett is taking the modality in different ways at different times, I will assume as a default notion of verifiability one that is exactly parallel to the notion of provability I worked out in Sect. 3.8. That is, I will say of a statement that it is verifiable iff it is decidable, and the decision method would lead to a verification of the statement. We might already have employed the method, or we might not. We do not even have to be sure which outcome applying the method would have, but we do need to know that applying the method would decide the issue.

---

[2] Together with the earlier chapter on intuitionistic logic, this chapter contains most of what I will have to say about purely verificationistic theories. Stage I, unlike the other stages, will thus not receive its own chapter in the third part.

As with the corresponding idea of provability, I believe that this makes for the most useful precisification of the notion of verifiability. It is, once again, a tensed notion. Statements that are not verifiable today may well become verifiable tomorrow. An untensed alternative would have to make recourse to an abstract realm of verifications that existed independently of our epistemic capacities, just as Prawitz's untensed provability had to employ a self-subsisting realm of proofs. If anything, the idea of such an independent realm of verifications seems even less plausible, given that empirical statements, unlike mathematical ones, include future tensed ones. To say that there might now be a verification of something that will happen in the distant future, whether or not we are able to devise a method of finding this verification, strikes me as too bold a proposition, and the quote back on p. 48 showed that Dummett would not have liked to make such a strong assumption either.

With this notion of verifiability, though, we run into a terminological problem that might lead to confusion: Although it fits Dummett's usage, it does not correspond to the notion of verifiability as it is understood in the writings of logical positivists. There, a statement is usually taken to be verifiable iff there are one or more observation sentences that together logically imply the statement. In other words, it is verifiable iff it is decidable, and the decision method could lead to a verification, though in fact it might also lead to a falsification. The root of this difference in usage is that the logical positivists wanted to use verifiability as a criterion of meaningfulness (and of course false statements should be considered meaningful as well as true ones), whereas Dummett wants to use verifiability as a criterion or surrogate for truth. Note that, interestingly, no similar problem seems to apply to our usage of "provable": No one I know of would want to say that "2 + 2 = 5" is provable.

While we are on the topic, let me note here that *falsifiability* will below likewise be interpreted along the same lines as verifiability and provability. We will consider a statement to be falsifiable iff it is decidable (i.e., we have a method of which we know that it will resolve the question), and the decision method leads to a falsification, whether or not anyone knows of that outcome. Again, care must be taken to avoid confusion with the notion of falsifiability that is often found in the philosophy of science, for example the work of K. Popper. There it will, as in the case of verifiability, not be required that the decision method should actually lead to a falsification.

In this work, an important role will be played by the idea of *unfalsifiability*. I will say that a statement is *pro tempore* unfalsifiable (or sometimes simply unfalsifiable) already if we, at the moment, lack a method of which we know that it will resolve the issue. This of course could change as we progress in our inquiry. I will assume that the fact that we (momentarily) lack knowledge of such a method is something we can recognize. On the other hand, I will say that a statement is recognizably eternally unfalsifiable in those cases in which it is known that a falsification will never be obtained.

Now, back to verificationism. In general, Dummett's verificationistic theory of meaning is of course a direct inheritance of Schlick's slogan "the meaning of a proposition is the method of its verification".[3] The verificationistic theory of meaning

[3] Schlick (1936), p. 341.

had been championed by him and the other members of the Vienna Circle, as well as by A.J. Ayer, who popularized the view in Britain.

Indeed, Dummett has much sympathy with those verificationists, though he of course differed with them on important points (apart from the terminological difference in use of "verifiable" just discussed). Dummett naturally disagreed with them when they claimed that all metaphysical statements were meaningless. And he thought that they would have come to agree with him, had they thought matters through, that on their account of meaning classical logic would have turned out to be unacceptable.[4] The emphasis on verifications should (just like the emphasis on proof in mathematics) have lead them to intuitionistic logic.

To show that intuitionistic logic is in fact the logic adequate for verificationistic theories of meaning, a piecemeal definition of the verification conditions of complex statements is called for. Recall how this was achieved in the case of intuitionistic mathematics. Here is the BHK interpretation again:

- c is a proof of $A \wedge B$ iff c is a pair $(c1, c2)$ such that $c1$ is a proof of A and $c2$ is a proof of B
- c is a proof of $A \vee B$ iff c is a pair $(i, c1)$ such that $i = 0$ and $c1$ is a proof of A or $i = 1$ and $c1$ is a proof of B
- c is a proof of $A \supset B$ iff c is a construction that converts each proof of A into a proof of B
- nothing is a proof of $\perp$
- c is a proof of $\sim A$ iff c is a construction which transforms each proof of A into a proof of $\perp$.

Alternatively, $\sim A$ was defined as $A \supset \perp$. $\perp$ is some mathematical absurdity such as $1 = 0$.

The most straightforward verificationistic adaptation of this interpretation would be this:

- c is a verification of $A \wedge B$ iff c is a pair $(c1, c2)$ such that $c1$ is a verification of A and $c2$ is a verification of B
- c is a verification of $A \vee B$ iff c is a pair $(i, c1)$ such that $i = 0$ and $c1$ is a verification of A or $i = 1$ and $c1$ is a verification of B
- c is a verification of $A \supset B$ iff c is a procedure that converts each verification of A into a verification of B
- nothing is a verification of $\perp$
- c is a verification of $\sim A$ iff c is a procedure that transforms each verification of A into a verification of $\perp$.

$\sim A$ can, once again, be defined as $A \supset \perp$ instead of giving the last clause explicitly. As we are here not talking about mathematics, it might not be suitable to take $\perp$ to abbreviate $0 = 1$ any more. It may instead have to be some *empirical* absurdity, such as "The moon is populated by pink flamingoes." We will get back to the meaning of $\perp$ soon.

---

[4] LBM, p. 10.

To tighten up the account so as to sustain a soundness and completeness proof, one would then simply turn to the Kripke semantics described in Sect. 3.7 and resolve to read value 1 as "verifiable" instead of "provable."[5] As mentioned above, it will come naturally to read verifiability along the same lines as I spelled out for provability in Sect. 3.8.

However, Dummett is not quite happy with such a simple translation from the mathematical to the empirical realm. He suggests that it is far from clear that the above clause for negated statements makes sense. Indeed, we might have to take recourse to falsifications to fix this problem.

5.3 Stage II: Expanded Verificationism

Here is the section that contains both Dummett's worry and his proposal to bring in falsifications:

> [A] proof of the negation of any arbitrary statement then consists of an effective method for transforming any proof of that statement into a proof of some false numerical equation. Such an explanation relies on the underlying presumption that, given a proof of a false numerical equation, we can construct a proof of any statement whatsoever. It is not obvious that, when we extend these conceptions to empirical statements, there exists any class of decidable atomic statements for which a similar presumption holds good; and it is therefore not obvious that we have, for the general case, any similar uniform way of explaining negation for arbitrary statements.
>
> It would therefore remain well within the spirit of a theory of meaning of this type that we should regard the meaning of each statement as being given by the simultaneous provision of a means for recognizing a verification of it and a means for recognizing a falsification of it, where the only general requirement is that these should be specified in such a way as to make it impossible for any statement to be both verified and falsified. (Dummett (1993), p. 71)

The worry thus is that no empirical statement, no matter how absurd it may be, has any arbitrary statement as a consequence, and the remedy Dummett proposes is to explain negation in terms of falsifications. Although I am in sympathy with this latter move, it has to be said that the worry is not yet grave enough to warrant it. I will show below how the argument can be strengthened sufficiently, but first let me tell you why Dummett's doubt does not yet push us irresistibly toward falsifications in the ingredient sense.

As we know, the intuitionistic account of negation involves two ideas: The first idea is that a verification (or proof) of a false statement can be transformed into a verification of an absurdity. The second idea is the one Dummett objects to, that a verification of an absurdity in turn can be transformed into a verification of any statement whatsoever. But we also know that if this idea and its corresponding rules

[5] Indeed, in his book *Elements of Intuitionism,* Dummett immediately gives "p is verified at w" as a gloss for $v_w(p) = 1$ (p. 139).

are dropped, we still get an account of negation. We end up with the negation of Johansson's minimal logic.[6]

To claim that no satisfying constructive negation can be defined by the above clause and thus to make a case for the introduction of falsifications, we would also have to object to the first idea, the transformation of a verification of a falsehood into a verification of $\bot$. And indeed, there is good reason to do so.

Just as it is not clear which statement might be so absurd as to be turned into a verification of any other statement, it is equally doubtful whether there is a statement such that any verification of a false statement would lead to a verification of it. To verify that my body height is under a meter (it is not) means, among other things, to apply a measuring tape to my frame. There is just about no way we should expect to come up with to turn that (non-actual) measuring episode into a verification of, say, the claim that the moon is populated by pink flamingoes. That means that this claim, absurd as it may be, cannot play the role of $0 = 1$ in the mathematical case.

What other statement could? The only candidates general enough seem to be statements like "A false statement has been verified": Indeed, if a false statement has been verified, then this instance of $\bot$ will quite trivially be true. However, how should the transformation process look like that morphs a verification of a falsity into a verification of this statement? It seems that the falsity of the mathematical statement that leads to a proof of "0 = 1" can be inferred from the absurdity of the latter. However, for "A false statement has been verified," we seem to have matters upside down. The only way to make the required inference in general will be to know of the statement's falsity beforehand. It is quite unlikely that this route will prove fruitful.

If we choose not to rely on absurdities that cannot deliver what we need from them, then we are indeed in need of some other means to explain negation. As is apparent from the quote above, for Dummett this is exactly the point at which *falsifications* enter the picture.

5.3.1 Intuitionistic Logic?

Something that is not quite as apparent from the quote above or the discussion surrounding it[7] is this: How *exactly* are we going to explain negations in terms of falsifications?

Here is what I believe to be both the most natural and, incidentally, the best way of doing it: We should simply say that a negated statement $\neg A$ is verified iff A is falsified.

This approach will require the availability of falsification conditions for both atomic and complex statements. In particular, we have to decide when a negated statement $\neg A$ is falsified. The natural answer is that $\neg A$ is falsified iff A is verified.

---

[6] See Sect. 3.6.1.

[7] Indeed, I do not know that Dummett addresses the question anywhere at all.

Negation then becomes a kind of toggle switch between verifications and falsifications. I will refer to such a negation as a *toggle negation*.

Toggle Negation:

$\neg A$ is verified iff A is falsified; $\neg A$ is falsified iff A is verified.

Once both verifications and falsifications are in play, it is very plausible that we should have a means of switching back and forth between them and that negation should be the device to do that. Toggle negations can occur in quite different logical systems. The exact behavior of a logic with toggle negation will depend on how we decide to fix the verification and falsification conditions of the other types of complex statements. I will get back to this task in the chapter on Stage II theories (Chap. 8); the falsification conditions I shall argue for will bring us to one of the so-called *Nelson logics*.

However, one thing is already clear at this point: No matter which falsification conditions for the other connectives we choose, we will not come out with intuitionistic logic. This is because every statement will be logically equivalent to its double negation. We have that $\neg\neg A$ is verified iff $\neg A$ is falsified iff A is verified. Also, of course, we have that $\neg\neg A$ is falsified iff $\neg A$ is verified iff A is falsified. As in intuitionistic logic, Double Negation Introduction (DNI) will be valid in a logic with a toggle negation. But *unlike* intuitionistic logic, Double Negation *Elimination* (DNE) will also be valid.

Now, Dummett does not mention the possibility that the introduction of falsifications at the level of ingredient sense would have any effect on the logic. So, we must presume that he thought that it would be possible to work falsifications into the original account somehow to motivate intuitionistic logic.

Of course, this might be right. Even if the above account of negation as a toggle between verification and falsification might strike us as natural, there still might be a way to stick with the old setup. Two questions present themselves: The first is "Why?," and the second is "How?"

First then, *why* would we want to preserve the original account of negation as an implication of absurdity? The intuitionists had to come up with their crafty way of defining negation because there seemed to be no other constructive way of accounting for negative information in mathematics. They only had recourse to the positive notion of *proof*.

Here, however, the situation is quite different. We already have introduced a negative primitive notion, that of a falsification. To eschew the simple toggle account above for a more intricate account seems to be a piece of needless ingenuity.

Dummett has lately admitted to having become sentimentally attached to intuitionistic logic.[8] Understandable as that may be, it makes for bad philosophical motivation.

Frankly, the only good reason I can see for turning to the intuitionistic account for a constructive empirical negation would be to argue that the initially less contrived idea of a toggle negation cannot, in the end, be expanded to a logical system that satisfies

[8] Auxier and Hahn (2007), p. 489.

constructivist demands. For example, a natural worry might be that a negation that satisfies DNE will automatically turn any constructive logic into classical logic. I will argue that this is not the case and that the Nelson logic that we will meet in Chap. 8 will do nicely for a constructivist. Moreover, it will turn out that intuitionistic negation is definable in Nelson logic (while toggle negation is not definable in intuitionistic logic). Therefore, if we see something in intuitionistic logic that we do not want to go without being able to express, Nelson logic can answer to that need as well.

All this will be dealt with later. For now, let us see how a determined intuitionist who, for whatever reasons, finds toggle negation objectionable, could go about setting up his negation in terms of verifications and falsifications. When we keep up the old account of negation as defined as an implication $A \supset \bot$, then there are two possible ways in which falsification conditions could add to the account: It could be the falsification conditions of A that make the account plausible, or the falsification conditions of $\bot$.

Let us start with the first of these options: We keep the usual interpretation of $\bot$, viz. $1 = 0$ or some empirical absurdity. In explaining how a verification of the false empirical statement A could be turned into a verification of $\bot$, we employ the new notion of a falsification and assume that A is falsified. Because A is falsified, we know that there can be no verification of it, and we will be able to claim that we can turn any such non-existent verification into a verification of $\bot$, because we will never be required to actually do it. We are dealing with what I called in Sect. 3.6.1 an *empty promise conversion*.

Now, note the following: This attempt tries to play on the fact that A is actually falsified (or falsifiable). It then seems that the verification clause for negation just amounts to the same as for toggle negation:

$\neg A$ is verified iff A is falsified.

Whether or not the account of the falsification of a negation will come to the same as for toggle negation (viz., $\neg A$ *is falsified iff* A *is verified*) as well will depend on what we would like to say about the falsification condition of a conditional (because we are taking $\neg A$ to be defined as $A \supset \bot$). Although this is the topic of a later section, I can already tell you that the most plausible account is this: A conditional is falsified iff its antecedent is verified and its consequent is falsified. If we assume that $\bot$ counts as falsified, then the only thing of substance that this condition will require is the verifiability of the antecedent. Thus, what we get is, once again,

$\neg A$ is falsified iff A is verified.

So this way of introducing falsifications into the account to come up with a better negation will just give us a more roundabout way of saying the same we said above, and therefore, it will not give us intuitionistic logic.

Let us then investigate the second option. The task here is to employ the notion of a falsification to come up with a suitable candidate for $\bot$.

If we insist that $\bot$ is one particular falsified statement, then it is clear that we have not made much progress. What before was called an *absurd* statement is now a *falsified* statement. At most, this might have helped if we had a candidate $\bot$, but were worried that we could not make sense of the notion of "absurdity." But no matter what we call it, our problem was that we did not *have* a candidate to fulfill the necessary inferential role in the first place.

But there is actually no reason why $\bot$ should have to be one particular statement. We might say that $\bot$ is simply *any* falsified statement. To verify a statement, $\neg A$ would then be to supply a method to convert every verification of A into a verification of some falsified statement. In this case, as above, we will have to give falsification conditions for all statements. Perhaps we can be a bit more economical here, though: We might be able to do with falsification conditions for *atomic* statements only and have $\bot$ denote some falsified atomic statement.

What we have here is, on the face of it, actually not all that implausible. To verify a statement of the form $\neg A$, find some falsified statement and show that a verification of A would lead to a verification of that falsified statement.

One problem here might be that no one can know which statement you took to be the relevant instance of $\bot$, unless you tell them explicitly. But that might not impede understanding much more than the original account of negation in intuitionistic mathematics did. There you did know which absurd statement was claimed to be derivable, but you did not know how that would be pulled off by the speaker unless she told you. Here, you do not know how the transformation is supposed to work, plus you do not know what the transformation will lead to; however, you know that it will be *some* falsified statement or other.

Even if we were required to know which falsified statement is the relevant instance of $\bot$, such a negation may also be expressed by making the underlying conditional quite explicit: "If the sun had exploded two days ago, the climate would have changed dramatically. But it has not changed. So the sun did not explode two days ago."

Here, a worry might arise from the following fact: The report that the statement playing the role of $\bot$ is actually falsified is effected by some means of negation ("The climate has not changed"). But negation is what we want to explain! Apparently, for this not to end in a vicious regress, there must be some other way of communicating that fact. For atomic statements, this might work by verbal or actual ostension: "Take a look out of the window!" This again suggests that it would be a good idea to take $\bot$ to be an atomic statement. Then it would be possible to give the meaning of negation in the following way: For atomic statements, the negation will be verified iff the atomic statement is falsified. For complex statements, the negation will be verified if a verification of the complex statement could be turned into a falsification of an atomic statement.

So far, maybe so good. At this point, we have to consider that any falsified atomic statement could be employed to play the part of $\bot$. Now, if we want to get full intuitionistic logic and not be stuck with minimal logic, we once again have to make plausible that every verification of any instance of $\bot$ could be turned into a verification of any statement whatsoever. But how could this be plausibly maintained? Would it really be possible to turn any verification of "The climate has changed dramatically," say a thorough look out of the window, into a verification of "Madrid is in Switzerland?" Again, the prospects of a *truly constructive conversion* look dim.

What the intuitionist has to say at this point, presumably, is this: It is not the atomic statement in itself that makes the transformation possible, but rather the fact that it is an instance of $\bot$ and hence falsified. As there are no verifications of falsified statements, the possibility of the transformation is vacuously satisfied: We can always claim to be able to turn something into something else if we are sure that we will never be challenged to execute this maneuver. Once again, it is an *empty promise conversion* that we have to employ.

So, to sum up this section, the intuitionist can stick to his account of negation if he is willing to accept *empty promise conversions* as constructively unobjectionable. However, as I said before, I will later present a more attractive account, built up around the notion of a toggle negation. This account will be natural and constructive, and intuitionistic negation will be definable in it. However, you will have to wait until Chap. 8 to see the details.

For now, it is enough to have a rough idea of how an expanded verificationism (a Stage II theory) might look. We have made room for falsifications in the ingredient sense of statements, but until now, they play no role in determining the assertoric content of a statement. If we allow this, then we enter the third stage of the taxonomy, the realm of the *hybrid strategies*.

5.4 Stage III: Hybrid Strategies

Dummett acknowledges that, given the need for falsifications in the ingredient sense, it would be quite natural to employ them in the assertoric content as well. However, he has grave reservations against this; indeed, he categorically rejects the idea of a hybrid strategy:

> It was conceded above that [a constructivist theory of meaning] may have to allow that what is taken to constitute the falsification of a statement must be separately stipulated for each form of sentence. But, if so, this can only be for the purpose of laying down the sense of the negation of each sentence, no uniform explanation of negation being available: it cannot be for the purpose of fixing the sense of a sentence, considered as being used on its own. (WTM, p. 76)

In terms of the distinction he adopted later, Dummett holds that falsifications may feature in the *ingredient sense* of a statement, which determines how it contributes to the sense of more complex statements. They may not, however, enter into its *assertoric content*, the sense that determines whether an assertion of it will be judged to have been correct or incorrect. We have to stay at Stage II and must not be tempted by a Stage III approach.

If both verifications and falsifications were to play a role in determining the assertoric sense of a statement, then they both would have to have an influence on our verdict whether an assertion of it was correct or incorrect. The only way that Dummett can imagine this could be made precise is to say that an assertion is correct iff it is verifiable and incorrect iff it is falsifiable. This is a natural thought, even though I will suggest two other forms a hybrid strategy can take. I will refer to the present strategy, in want of a catchier title, the *correctness as verifiability and incorrectness as falsifiability* strategy, and abbreviate this from now on as CV&IF.

This strategy would lead to a problem in that, according to Dummett, it would allow for gaps between correct and incorrect assertibility. Such gaps, however, are not to be admitted:

> It would then follow that a speaker might be neither right nor wrong in making an assertion: not wrong, because it could be shown that the sentence could not be falsified; but not right either, because no way was known of verifying the sentence. This consequence would be fatal to the account, since an assertion is not an act which admits of an intermediate outcome; if an assertion is not correct, it is incorrect. (WTM, p. 77)

Dummett stresses that this refusal to make room for assertions that are neither correct nor incorrect is not supposed to mean that we have to be able to determine of each assertion whether it is correct or incorrect. He elaborates this point thus:

> If the content of an assertion is specific, then it must be determinate, for any recognizable state of affairs, whether or not that state of affairs shows the assertion to have been correct. If some recognizable state of affairs does not suffice to show the assertion to have been correct, there are two alternative cases. One is that this state of affairs serves to rule out the possibility of a situation's coming about in which the assertion can be recognized as having been correct: in this case, the state of affairs must be taken as showing the assertion to have been incorrect. The other is that the given state of affairs, while not showing the assertion to have been correct, does not rule out the possibility of its later being shown to have been so: in this case, the correctness of the assertion has simply not yet been determined. What is not possible is that any recognizable state of affairs could serve to show both that the assertion was not correct and that it was not incorrect, since the content of the assertion is wholly determined by which recognizable states of affairs count as establishing it as correct: so any state of affairs which can be recognized as ruling out the correctness of the assertion must be reckoned as showing it to be incorrect. Hence, if a sentence is held to be neither true nor false in certain recognizable circumstances, this cannot be explained by saying that an assertion made by uttering the sentence would, in those circumstances, be neither correct nor incorrect. (WTM, p. 80)

Thus what Dummett objects to is the possibility that we may come to know that an assertion is neither correct nor incorrect. He is not claiming that we will always know whether it was one or the other.[9]

It is not a straightforward matter how to make sense of this. The pattern is reminiscent of the double bind the intuitionists find themselves in due to their allegiance to tertium non datur and their rejection of bivalence. This peculiar situation, as you will recall, can be seen as emanating from the validity of LET ($\sim\sim(A \vee \sim A)$) and the invalidity of LEM ($A \vee \sim A$) in intuitionistic logic. In fact, one way of grounding the above contention that "any state of affairs which can be recognized as ruling out the correctness of the assertion must be reckoned as showing it to be incorrect" would be to argue that an assertion is incorrect iff (1) it would have been correct to assert its negation and (2) that negation has to be interpreted intuitionistically.

---

[9] Cf. also WTM, p. 78: "[T]here cannot be a piece of knowledge the possession of which by any speaker would show both that he would not be right to make a certain assertion and that he would not be wrong to make it."

However, neither of these claims is trivially true. Regarding assumption (2): I have argued briefly above that, once we are ready to introduce falsifications into the ingredient sense, it is not perfectly clear any more that intuitionistic negation is the only, or even the most attractive, game in town. Consequently, it is not clear that we have to follow the intuitionists into every difficult situation that is created by their account of negation. For example, if we adopt a toggle negation account, it is clear that LEM and LET will stand or fall together, depending on how the rest of the theory is spelled out.

No, the argument that we cannot come to know of gaps between correct and incorrect assertibility must come from a further analysis of these notions. One part of this analysis will have to settle whether assumption (1) is justified. I will turn to this task in the next chapter.

But even taking it, for now, as given that an assertion cannot be known to be neither correct nor incorrect, it is not easy to understand what Dummett is driving at in those two quotes. They are from the same essay, but reading them in sequence is puzzling, as it is hard to keep in view where the problem is supposed to lie.

Consider the case envisaged in the first quote: A certain assertion is "not wrong, because it could be shown that the sentence could not be falsified; but not right either, because no way was known of verifying the sentence."

However, this assertion would, if correctness and verifications go together as indicated in the second quote, be described thus: We would know that the assertion is not incorrect, because we know that there can be no falsification of it. But whether it is correct is, as it has not been verified, "simply not yet determined." The situation is thus quite distinct from one in which we know that the statement is neither correct nor incorrect.

This case, of course, *could* arise as well (all this under the assumption that assertions are correct iff the asserted sentence is verifiable and incorrect iff it is falsifiable). Such would be a case in which we know that the sentence will never be verified and that it will never be falsified. Contrary to what one may think on first blush, this is a situation that also the intuitionist can envisage, even if it would be quite a weird situation from her point of view.

It would, for them, not constitute a breach of the LET, because the information that the statement could not be verified would already warrant assertion of the negation of the statement. Thus, if we were to tie incorrectness to falsifiability, an assertion of the statement would not be incorrect, but an assertion of its intuitionistic negation would be correct. This is because the intuitionistic negation requires only that the negated statement can never be verified (because its verification would lead to something absurd), not that it can be falsified.[10]

Intuitively, this seems indeed like a strange combination; but again, it presupposes the intuitionistic account of negation. Therefore the strangeness of this combination can only count against statements that are unverifiable and unfalsifiable in principle and the ensuing assertibility gaps if this intuitionistic presupposition is upheld.

[10] In Sect. 6.4, I will consider and reject the idea that a verification of the impossibility to verify a statement should already count as a falsification of it.

If, instead, we employ the toggle negation, we get that neither the statement nor its negation is correctly assertible, and neither would assertions of them be incorrect. Again, that such an account is incoherent must be argued for by spelling out in more detail what it is to make a correct and what it is to make an incorrect assertion. And once again, I will turn to that question in the next chapter.

If we assume that assertibility gaps are not yet conclusively ruled out, then a hybrid strategy that takes correctness to be correlated to verifiability and incorrectness to falsifiability is likewise not yet off the table. However, besides such a CV&IF theory, there are two other kinds of theories at Stage III that I would like to offer up for consideration. Both of these accept the Dummettian contention that there are no assertibility gaps, and both are strategies to manage two *different* accounts of what it is to make a correct assertion.

These different accounts lie at the respective hearts of the ideas I have simply called *verificationism* and *falsificationism*. The verificationist claims, as we have seen, that a correct assertion is one that can be verified. The falsificationist, in contrast, holds that a correct assertion is one which cannot be falsified.

Before I sketch the ways, I believe these accounts can be combined at Stage III, let me tell you about falsificationism on its own. The theories that are *only* governed by the falsificationistic norm of assertion are to be found at the remaining two stages, Stage IV and Stage V.

5.5 Stages IV and V: Expanded and Pure Falsificationism

Now, why should we want to adopt a falsificationistic theory? This question is the main theme of the next chapter, but I will give a first glance in the present section.

The idea of a falsificationistic norm of assertion is Dummett's own. However, as I mentioned in the first section of this chapter, Stage IV is not among the options Dummett considers. Instead, after proclaiming Stage III to be untenable, he devotes several pages to outlining a Stage V theory. That is, a meaning theory that has no need at all for verifications, whether in the assertoric content or the ingredient sense.

He also adds a short characterization of a logic that would be motivated by such a semantic theory, a logic that I will examine in detail later on. The need to investigate an expanded falsificationism, that is a Stage IV theory, will arise from problems we will find with the purely falsificationistic theory and the meanings that Dummett thinks should be given to the logical constants in such a theory. As one might expect, these are quite similar to the problems with pure verificationism we have found in the first part of this chapter.

Now, I said that the central idea of falsificationism is that an assertion is correct iff it is unfalsifiable. However, it might be even better expressed as follows: An assertion is *in*correct iff it can be falsified. A falsificationistic theory of meaning is, according to Dummett, motivated by the following realization: The incorrectness of an assertion is prior to its correctness in our understanding of it.

> The fundamental notion for the account of the linguistic act of assertion is (...) that of the incorrectness of the assertion (...) By making an assertion, a speaker rules out certain possibilities; if the assertion is unambiguous, it must be clear which states of affairs he is ruling out and which he is not. (...) Thus, in the order of explanation, the notion of the incorrectness of an assertion is prior to that of its correctness. (WTM, p. 124)

The idea that assertions guide the audience by excluding possibilities can be found elsewhere as well. Take one extremely basic example as an illustration, adapted (with a slight modification) from van Benthem[11]:

A family eats out in a restaurant, and a waiter comes to take their orders. The father asks for a meat dish, the mother for a vegetarian one, and the child for fish. After a while, a different waiter comes out and brings the plates. Seeing the slightly puzzled look on the new waiter's face, the family realizes that he has no idea which of the six combinatoric possibilities of handing out the dishes is the right one. Helpfully, the father says, "The meat is for me," thus reducing the possibilities to only two. The mother goes on to exclude one more option by saying, "I asked for the vegetables," and the two assertions together have reduced the set of possibilities so successfully that the child does not have to say anything more to have the waiter distribute the plates correctly.

This process of weeding out wrong possibilities is of course also what Popper claims drives science. Furthermore, it is the idea that lies at the heart of epistemic modal logic, where the possibilities are represented as worlds in Kripke models.

In Dummett's theorizing, this simple idea is further developed as follows: On constructivist assumptions, we have to be able to recognize the states of affairs that define the content of an assertion in order to understand it. And these, according to the present proposal, are the ones that are excluded by the statement. In other words, what we have to grasp are their *falsification* conditions.

> These considerations prompt the construction of a different theory of meaning, one which agrees with the verificationist theory in making use only of effective rather than transcendental notions, but which replaces verification by falsification as the central notion of the theory: we know the meaning of a sentence when we know how to recognize that it has been falsified. Such a theory of meaning will yield a logic which is neither classical nor intuitionistic. (WTM, p. 83)

The assumption that what a statement excludes is exactly delineated by the states of affairs that falsify it is made by Dummett without further comment. Basically, the same general line of thought as above can already be seen in his early essay "Truth" (1959), even though there is no explicit mention of falsifications yet:

> A statement, so long as it is not ambiguous or vague, divides all possible states of affairs into just *two* classes. For a given state of affairs, either the statement is used in such a way that a man who asserted it but envisaged that state of affairs as a possibility would be held to have spoken misleadingly, or the assertion of the statement would not be taken as expressing the speaker's exclusion of that possibility. If a state of affairs of the first kind obtains, the statement is false; if all actual states of affairs are of the second kind, it is true. (TOE, p. 8)

---

[11] Cf. van Benthem (2008).

Here, it would be important to know what exactly it means that a state of affairs obtains. Once again, there is no way to be completely sure how Dummett meant this, but to me it seems most plausible to say at this point that the obtaining states of affairs are the states of knowledge we might obtain, given our currently available methods.

Now, it is one thing to say that the content of a statement is clearly given by what it seeks to exclude. It is quite another to say that a statement is correct (or even *true*) if nothing it claims to exclude can be established. How well this actually fits our intuitions will, I think, depend on the area of discourse we are looking at.

To give an example in which the proposal looks quite implausible, take once again mathematics. Of course I can assert Goldbach's conjecture, and it will be quite clear what I seek to exclude by my asserting it. But even though no counter example can be given at the present moment, I will not be looked at as having made a correct mathematical assertion. If anything, the new way of judging a mathematical speech act correct will apply not to assertions, but to conjectures. These will stand as long as they are not falsified and will be elevated from their status to that of theorems once a proof is found.

After some more comments on the nature of correctness and incorrectness, I will spend the rest of the next chapter by suggesting some areas of discourse in which, I believe, the proposal has a better chance of being accepted. The most promising example will be taste talk, as in "Spinach is tasty."

5.6 Hybrid Strategies Again

This brings us back to the question which form a hybrid strategy can take. Assume that indeed for some areas of discourse verificationism delivers a better model, while in others falsificationism appears more attractive. In that case, a rather trivial form of a hybrid strategy suggests itself: Employ verification conditions to determine the assertoric sense in those areas in which verificationism seems more plausible and let falsification conditions play the lead role in those areas in which falsificationism reigns.[12]

I will call such a strategy a *discourse separation* strategy. Unlike the other hybrid strategy we have seen above, the CV&IF approach, a discourse separation strategy, will not give rise to gaps between correct and incorrect assertibility. In the verificationistic areas, an assertion will be correct iff verifiable, and incorrect otherwise. In

[12] When I say that this is trivial, I mean that it is easy to explain how verifications and falsifications are related to each other in the assertoric sense. The question of whether and how we can clearly separate different areas of discourse, on the other hand, is not trivial at all. Is "The number of tasty dishes on this menu is prime" a statement from the area of taste talk, or from the mathematical area? It is frequently assumed that a separation of areas of discourse is feasible, but it is actually far from clear to me how it can work in detail. However, I will not go into this problem in any depth in this book.

the falsificationistic areas, an assertion is incorrect iff falsifiable, and correct otherwise.

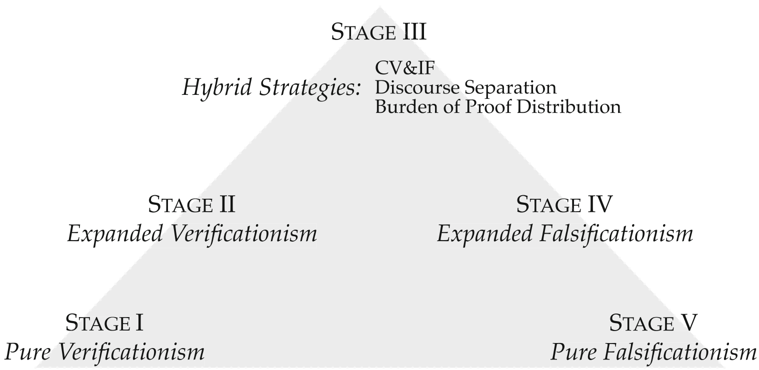
The third option I'll propose will likewise respect Dummett's disdain of such gaps, but it will strive to combine the verificationistic and the falsificationistic account of assertibility *in the same area of discourse*. The crucial device will be the notion of a *burden of proof*, a burden that may either lie on the speaker or on the audience. If the speaker bears the burden, assertibility will coincide with verifiability, as it is more commonly assumed. If the hearers have to bear it, then the assertibility will be construed as unfalsifiability. I will call this the *burden of proof distribution* strategy, or simply the *burden of proof* strategy.

I will supply more details about the exact ways a hybrid strategy can be set up in the chapter about Stage III.

5.7 Chapter Summary

I have discerned five different stages of falsificationistic involvement. Four of these are explicitly mentioned by Dummett. He started out with a pure verificationism (Stage I), which he had to expand to allow for a more plausible account of negation (Stage II). He dismissed, however, the idea that falsifications and verifications might both determine the assertoric content (Stage III). Instead, he toyed with the idea that this assertoric content might be determined by falsifications only and that verifications might be dispensed with completely (Stage V).

The stage not considered by Dummett (Stage IV) suggests itself once we become aware of problems involving the meaning of complex statements in the purely falsificationistic theory (Stage V). Those problems are very similar to those in the purely verificationistic theory (Stage I) that lead to the introduction of falsifications in the first place, and the remedy is also similar: The purely falsificationistic theory will be augmented by verifications in the ingredient sense, resulting in an expanded falsificationism.



Here we see the pyramid again. This time I have added the three options that I see for someone who, against Dummett's advice, opts for a hybrid strategy (Stage III).

First, we saw the variant Dummett had in mind and argued against. This is the *correctness as verifiability and incorrectness as falsifiability* strategy, CV&IF for short.

The other two options, in contrast, combine verificationism and falsificationism.

Firstly, I talked about the *discourse separation* strategy, which simply entailed that different areas of discourse were governed by different norms of assertion.

Secondly, I mentioned the *burden of proof* strategy. Here, both verificationism and falsificationism govern the same area of discourse. Under which of these norms an assertion will be judged to be correct or incorrect will depend on where the burden of proof lies at the moment of utterance.

Note that there is nothing that obviously forbids one to combine the *discourse separation* strategy with the other two: It might conceivably turn out that one area of discourse supports verificationism, a second one falsificationism, a third one a CV&IF theory, and a fourth one a *burden of proof* theory.[13]

Both of the latter hybrid strategies are basically mechanisms to switch from verificationism to falsificationism and back. Consequently, even if one of these hybrid strategies will turn out to be more attractive than what we find at the lower levels, we will need a good understanding of the verificationistic (I and II) and the falsificationistic (IV and V) theories. When I begin to discuss the stages one by one, it is therefore expedient to ascend the pyramid starting from the bottom.

The main concern of this step by step examination will be to discern the different logical systems that best fit the respective stages. Dummett himself only considered two different systems. He thought that both Stage I and II would motivate intuitionistic logic and that Stage V leads to a logic known as dual intuitionistic logic. (Stage III was rejected by Dummett as incoherent and Stage IV, once again, not considered.)

In contrast, I will suggest different logical systems for each stage. In other words: If Dummett is right in supposing that falsifications have some role to play in a correct theory of meaning, then no matter how great or small that role will turn out to be, intuitionistic logic will not necessarily be motivated by these considerations.

Before I turn to the logical details in Part Three, I will spend the next chapter on the idea that is common to all theories on the right side of the pyramid: To make a correct assertion is to say something unfalsifiable.

[13] Not that I believe such an extremely eclectic outcome likely.

Chapter 6
Falsificationism

6.1 Chapter Overview

In this chapter, I will investigate the central idea of the falsificationistic theories: An assertion is correct iff it is unfalsifiable.

I will go into Dummett's arguments for this thesis and try to disentangle what he has to say about the centrality of incorrectness, the defining character of our retraction behavior, and the possibility that truth itself might be nothing more than unfalsifiability.

To get a better grip on these very slippery issues, I will introduce some areas of discourse where the approach seems particularly appropriate. I will talk about disputes at criminal trials, where the defendant usually gets away with making assertions that are not falsifiable. I will also talk about *faultless disagreement*, a phenomenon that occurs in discussions about personal taste. As the name implies, faultless disagreement describes the semantically puzzling fact that there appear to be real disagreements about taste, but often neither party will be obviously at fault. I'll discuss this topic in greater depth, because I think it is the best chance for falsificationism to appear reasonable. Also, in this section, the doctrine of analetheism (see Sect. 4.12) will make a reappearance.

Coming back from the examples, we will see whether or not we have a better idea of what it is for an assertion to be correct, incorrect, true, or false. In the end, further spurred on by reflections on the indicative conditional, we will see Dummett suggest that the only property we need to grasp clearly in order to understand indicative conditionals is *recognizable incorrectness*. We will see that this is enough to build a notion of logical consequence on, so that at the end of this chapter, we will be ready to dive into the logical explorations of Part Three.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_6,

6.2 The Centrality of Incorrectness

In the last chapter, we saw Dummett propose, both in TRUTH and in WTM, the following central view: The content of a statement is determined by the states of affairs it rules out, and it is only false or incorrect if one of these states of affairs obtains. The positive notions of truth and correctness are derivative of the negative ones.

Let us see some more arguments for this emphasis on the negative notions. In WTM, Dummett writes:

> An assertion is not, normally, like an answer in a quiz program; the speaker gets no prize for being right. It is, primarily, a guide to action on the part of the hearers (...); a guide which operates by inducing in them certain expectations. And the content of an expectation is determined by what will surprise us; that is, by what it is that is not in accord with the expectation rather than by what corroborates it. The expectation formed by someone who accepts an assertion is not, in the first place, characterized by his supposing that one of those recognizable states of affairs which render the assertion correct will come to obtain; for in the general case there is no bound upon the length of time which may elapse before the assertion is shown to have been correct, and then such a supposition will have, by itself, no substance. It is, rather, to be characterized by his not allowing for the occurrence of any state of affairs which would show the assertion to have been incorrect; a negative expectation of this kind has substance, for it can be disappointed (WTM, p. 82).

But why is it that a disappointed expectation should have more substance than one that is fulfilled? The reason Dummett gives us is that the disappointment of such an expectation has a *systematic effect* in our dealings with assertions:

> There is a well-defined consequence of an assertion's proving incorrect, namely that the speaker must withdraw it (...) there is not in the same way a well-defined consequence of an assertion's proving correct (TOE, p. 20).

The idea is that an assertion that comes to be seen to have been correct might have an effect, or it might not. Only in certain circumstances like the mentioned quiz show will a correct statement elicit a predictable response. In general, though, while a correct assertion might have a positive effect, there is no guarantee for that. The only guaranteed feature of an assertion is that it will have to be withdrawn when shown wrong.

Here, Dummett sees a close analogy to the conventions governing commands. Obedience might or might not be rewarded, but the obedient has no *right* to expect such a reward. Disobedience, however, results in the right to reprimand the disobedient in one way or another. Therefore, it is central to our understanding of a given command to know what would count as disobedience; whether or not we are clear about which behavior should be called "obedient" is only of secondary importance.[1] If someone complied with a command unintentionally, did he obey or did he not? No matter that we might be confused about how best to describe his behavior; it is clear that no rebuke is in order, and knowing this shows that we have understood the command.

---

[1] See TOE, p. 8 and WTM, p. 82.

Likewise, Dummett holds that to understand an assertion is to be able to recognize the states of affairs it excludes. These states of affairs determine the content of the assertion because they oblige the speaker to retract it, and this is the only well-defined aspect of our assertoric practice.[2]

But is that really true? Aren't there any other aspects of our assertoric practice that are as systematic as our obligation to retract incorrect claims?

In fact, Dummett later came very close to saying that there *are* further aspects of this kind. Consider the following quote:

> It is fundamental (...) that an assertion may be judged as correct or incorrect by a hearer; the speaker may subsequently be compelled to withdraw it as incorrect, or the hearer to acknowledge it as correct. (...) The description of the practice of assertion will, among other things, delineate the situation in which a speaker is compelled to withdraw an assertion as incorrect, and that in which a hearer is compelled to acknowledge it as correct (LBM, p. 165).

This seems to contradict his earlier claim we just read, viz, that "there is not in the same way a well-defined consequence of an assertion's proving correct." There is at least one such consequence, namely that the audience has to accept the claim.

Is there any way to argue for the earlier position, for the view that the only feature we should focus on are retractions of statements? One *might* say the following: The fundamental norms governing language are *speaker norms.* Moreover, these norms are such that they only have practical consequences for a speaker who flouts them. On the other hand, what the audience does or refrains from doing is secondary. Furthermore, whether the hearers accept something or not may often be very hard to tell from their behavior, so there might be a further worry about manifestation of understanding here.

I'm not at all sure how water tight such a line of argument can be made. The main ideas of the falsificationistic theories of meaning, however, seem to depend on the assumption that incorrectness and retraction play the lead role in the determination of meaning. In my discussion of the falsificationistic theories, I will thus make this assumption of Dummett's and pick up the question how plausible it is when I get to the CV&IF strategy in Chap. 10.

---

[2] Dummett says that this has not dawned on people earlier because they were only thinking about a limited number of examples:

> Why has [the primacy of incorrectness] been so persistently overlooked? Partly because of the tendency to concentrate on the decidable case: an expectation as to the outcome of a test may indifferently be described as an expectation that the result will be favorable or as an expectation that it will not be unfavorable. Even, perhaps, in part because of a tendency to think particularly of future-tense assertions which predict the occurrence of an observable state of affairs within or at a specified time; for then the positive expectation has a bound, and hence has substance—if it is not satisfied within the given time, it will be disappointed (WTM, p. 83).

6.3 Retracting, Surprising, and Misleading

But even assuming (for now) that retraction behavior is the only feature we should base our theory of meaning on, I have some further remarks to make. First, I believe that the incurred obligation to retract assertions is not quite as strong as Dummett seems to suggest in the quote above ("... namely that the speaker must withdraw it").

Certainly, not all assertions that are later proven incorrect will have the result that the speaker withdraws it, and it would probably be too much to say that this already constitutes a breach of the conventions that govern our linguistic dealings. Suppose that yesterday I said to you that it would rain today and it turned out to be a sunny day after all. Am I really under an obligation to seek you out and retract my assertion, provided that nothing much was hanging on the question of today's weather? I would not want to say that.

However, this much is true: If a state that my assertion ruled out obtains and comes to your attention, then you have the right to request of me that I recant. This seems to me to be the core of Dummett's observation.[3]

On the other hand, I do not think that characterizing the states of affairs that a statement excludes by what would *surprise* the audience, nor by saying that "a man who asserted it but envisaged [such a] state of affairs as a possibility would be held to have spoken misleadingly" is a good idea.[4] The reason is that there are cases of such misleadingness that neither give rise to the obligation to retract the offending statement nor seem to have any connection to falsifiability.

To employ yet another classic example, let us suppose that you read in a letter of recommendation nothing more than that the candidate has very nice handwriting. Let us further assume that the job in question requires many skills and qualities, but that nice handwriting is clearly not among them. You should then be very surprised indeed to learn that the author of the letter actually thought that the candidate was splendidly qualified for the job and might even accuse the author of misleading you into thinking that the candidate was worthless.

Saying that the case in which the candidate was good was excluded by the assertion because in such a state the statement would have been misleading will result in a quite different account than saying that the states of affairs that are excluded by it are those in which the speaker is under an obligation to withdraw the assertion. The author of the letter, even when challenged, will be under no such obligation at all (assuming that it cannot be further pointed out that the candidate, highly qualified as she may be, has lousy handwriting, i.e., assuming that the statement cannot be falsified). He might have to give an explication why he made such a misleading assertion, but he will not have to abjure his claim that the handwriting was exquisite.

---

[3] I do not think Dummett would object to the thought in the last paragraph, and maybe even meant "... has to be withdrawn *when challenged*" to be understood implicitly. Sometimes, I will likewise leave off this qualification in what follows.

[4] TOE, p. 8.

If we want to follow Dummett in his singling out the states of affairs in which a statement has to be retracted as that feature which determines its content, we should hesitate in identifying these states with the states in which the statement would have been misleading.[5]

6.4 What is a Falsification?

Taking stock of the findings so far: According to Dummett (or at least according to his earlier view), an assertion's content is completely characterized by the states of affairs that are excluded by it. If such a state of affairs is pointed out to obtain, then the speaker will be under an indelible obligation to retract the assertion if challenged. Nothing so far is forcing us to think that the excluded states of affairs should have to be exactly those in which the statement is *falsified*.

Other options seem to be available, maybe even commendable. For example, we might want to say that what an assertion excludes are the states of affairs in which the speaker cannot verify (or cite a method of verification for) his claim, and thus stick to a form of verificationism after all. As I briefly mentioned in the last chapter, this seems to be a better way to view mathematical discourse. To make a mathematician retract an assertion (as opposed to a conjecture) of a mathematical statement, it is not necessary to falsify the statement. It is enough to point out that the proof he offered does not in fact establish the claim he makes.

Maybe, this worry can be dealt with by choosing a sufficiently broad notion of falsification. After all, we have not yet seen what exactly a falsification consists in. Indeed, Dummett himself does not tell us precisely. Maybe, pointing out a fault in an alleged proof should count as a falsification of a mathematical assertion.[6]

If we want to respect normal usage, however, we would not want to call this a falsification. It is clear what you have to do to falsify Goldbach's conjecture: Find an even number and show that there are no two primes that add up to it. Likewise, in empirical context, we think of a crucial experiment that rules out an accepted hypothesis, or, more mundanely, pointing out the empty mat to falsify the statement "The cat is on the mat." We would surely not interpret the information that a given statement has not been verified as a falsification of it.

Importantly, even the information that a statement *can never* be verified would by itself not pass as a falsification. For example, not every statement about occurrences that lie beyond the event horizon are automatically falsified. Maybe, we could capture that intuition by saying that falsifications are not concerned with our epistemic state

---

[5] My assessment here is largely in accord with Rumfitt (2007).

[6] Dummett speaks both of *falsifying a statement* and *falsifying an assertion*. The only sensible way to understand the latter is to read it as elliptical for *falsifying the asserted statement*. There are other instances of this elliptical use of *assertion,* and at times, I will use it this way as well to avoid unwieldy sentences; an assertion proper, though, is quite explicitly a speech act.

or our recognitional capacities.[7] That is not to say, of course, that falsifications are not *constrained* by our epistemic limitations, it is only to say that to falsify, a statement requires more than just pointing out that the verification of the statement lies beyond those limits. That additional element obviously ensures that a future verification is impossible, but this impossibility must lie in the nature of the state of affairs the statement describes, not in its epistemic accessibility. Thus, if a future verification of "There is no cow in the room (at time t)" is found to be impossible because we check and find a cow in the room, then *that* impossibility suffices to falsify the statement. But verifications that ensure this kind of impossibility seem to me to coincide exactly with what I think our intuitive notion of falsification amounts to anyway.

Nothing Dummett explicitly says about falsifications suggests that he has a notion that differs from this intuitive one in mind. Surely, he is not endorsing the first view above that a proof that something is presently not verified is sufficient for a falsification.[8] But then, at least for mathematical assertions, Dummett's identification of incorrect and falsifiable assertions does not seem very convincing. The idea that an assertion is incorrect iff it can be falsified can thus be questioned, at least in some areas of discourse, such as mathematics; I will below offer some example areas of discourse that seem more promising.

6.5 Correct Assertibility as Unfalsifiability

Let us for now turn to what Dummett says about the positive notions of correctness and truth. As we have seen, he claims that those assertions that cannot be falsified should be considered to be correct. This idea is, again, based on the assumption that correctness is derivative of incorrectness.

There are two similar claims to this effect, one in TRUTH, the other in WTM. Here is the quote from TRUTH again:

> A statement, so long as it is not ambiguous or vague, divides all possible states of affairs into just *two* classes. For a given state of affairs, either the statement is used in such a way that a man who asserted it but envisaged that state of affairs as a possibility would be held to have spoken misleadingly, or the assertion of the statement would not be taken as expressing the speaker's exclusion of that possibility. If a state of affairs of the first kind obtains, the statement is false; if all actual states of affairs are of the second kind, it is true (TOE, p. 8).

And here is the relevant quote from WTM:

> The fundamental notion for an account of the linguistic act of assertion is (...) that of the incorrectness of an assertion: the notion of its correctness is derivative from that of its incorrectness, in that an assertion is to be judged correct whenever something happens which precludes the occurrence of a state of affairs showing it to be incorrect (WTM, p. 82).

---

[7] Unless, of course, the statement to be falsified is about those capacities.

[8] For one thing, his worries about statements that are not verified (or even known to be unverifiable) and known to be unfalsifiable that were mentioned in the last chapter should not arise. There are no such statements if a proof that there is no verification is already a falsification.

One of the salient differences between the two versions is that in TRUTH, Dummett talks about the truth and falsity of a statement, whereas in WTM, he speaks of the (in)correctness of an assertion. Thus, these quotes seem to illustrate well the two modes of thinking that were discussed in Sect. 2.9, the one taking truth as the primary notion, the other assertibility. However, the way Dummett himself later interpreted what he wrote in TRUTH is that this essay also took assertibility as the primary notion.[9]

A more significant difference is the exact relationship between the positive and the negative notions. In WTM, the condition that is necessary for a positive evaluation is more restrictive than in TRUTH. It is not any more the mere absence of any states of affairs excluded by the statement, but the obtaining of something which *precludes* future falsification (At least, this is what I find the quote seems to be saying. It might also be possible to read it along the lines of the first quote, by reading "precluding the occurrence of a state of affairs" simply as determining that the currently available methods will not inform us that the state of affairs in question obtains).

In terms of correctness, which of the ways of determining it would seem more plausible? Would we want to call an assertion correct as long as nothing it excludes can be established, as TRUTH suggested? Or would we want to refrain from doing so until it was established that the statement could no longer be falsified, as the line ran in WTM? We are talking about unfalsifiability, but are we talking about *pro tempore* unfalsifiability or recognizably eternal unfalsifiability?

Up to now, our discussion is ailing from a lack of plausible examples. It seems clear that for mathematical discourse, none of the options seem particularly appealing if we are hoping to retain some of the intuitive sense we attach to "correctness." How about empirical talk, then?

Alas, it is equally questionable that typical empirical discourse would fare much better as an illustrating example. Will we judge assertions of an historian to be correct just because we cannot falsify his claims? Would he stand a better chance if he could adduce evidence that we could never settle the issue he addressed? While it is true that conclusive evidence is often not to be found in the discipline of history, it seems that the historian has to deliver a bit more.[10] And this is even clearer in cases where evidence is usually available. The zoologist who claims to have found a new species of jellyfish cannot rest his case on the observation that it will be impossible to prove that they do not exist.

---

[9] TOE, p. xxii.

[10] Else it could hardly be explained that "The Jew of Linz" (Cornish 1998) has gained infamy instead of acclaim as a historical work. In the book, it is argued that Wittgenstein and Hitler met when they were school boys; Hitler, it is claimed, was so annoyed by Wittgenstein (possibly after a short homoerotic infatuation) that later in life, he was unable to shake off his dislike and projected it onto all jews, with disastrous consequences. Cornish does claim that Hitler purged the records of the school in question and thus makes it quite plausible that counter evidence to his thesis will never be found. The usually disdainful verdict of the reviewers of the book (see for example Monk (1998)) is based on the fact that there is precious little evidence for Cornish's claims. They did not suggest that they could disprove those claims.

It is very unlikely, then, that unfalsifiability, whether *pro tempore* or recognizably eternal, should serve as the hallmark of correct assertibility in general, even if we exclude mathematics. If it is possible to give examples that plausibly motivate what we find at STAGE IV or STAGE V, it will presumably be by compartmentalizing and singling out specific areas of discourse, so that, in the end, we seem to be moving toward a STAGE III *discourse separation* strategy.

Before I return to the question which notion of unfalsifiability should amount to correctness, I shall try to point out some such areas. I will concentrate on two examples, legal and taste talk. I'll suggest that there is a clear sense in which defendants have to be taken to make correct assertions until they are falsified. I will discuss taste talk and other areas of discourse in which it seems possible to have *faultless disagreements*. A faultless disagreement is a situation in which two speakers are, on the face of it, defending incompatible positions ("Spinach is tasty" vs. "Spinach is not tasty"), but neither of them seems to be making a mistake.

6.6 Example I: Legal Trials

The first area of discourse I'd like to mention is legal talk, especially the assertions and claims one encounters in criminal trials. I do not mean claims about laws and their interpretation, but rather the claims prosecution and defense make about the events that occasioned the trial.

Any legal system that values the protection of the innocent higher than the punishment of the guilty is operating under the *presumption of innocence*: The innocence of the defendant is assumed at the outset of the trial and has to be disproven by the prosecution. This principle reflects an imbalance in our judgements about the severity of the outcome of possible errors. It is worse to convict an innocent man than to let a guilty one go free.

The presumption of innocence has an important effect on the discussions that take place at court: The prosecution is bearing the *burden of proof*. The prosecution has to prove its allegations, often "beyond reasonable doubt."

This indeed seems to make it plausible that what the accused says in his or her defense will be judged to be correct unless it is falsified. If such a falsification has not been delivered at the time of the verdict, the claim in question will stand. Of course, if they are later falsified, they will be judged incorrect then, but normally this has no influence on the verdict any more. Moreover, the judgement of having made an incorrect assertion will only be given for those things the defendant said that were falsified during the trial, not those for which he could not prove that a falsification would be impossible. Here, then, correctness is correlated with unfalsifiability *pro tempore*.

Indeed, I think that this setting is quite promising for the falsificationistic approach. However, it is clear from the outset that not *every* assertion made at a trial will be judged correct unless falsified. For one thing, the prosecution is in a position more reminiscent of the mathematician's in that her claims will only be accepted as correct if they can be verified. Such is the nature of the asymmetry that the presumption of innocence and the burden of proof produce.

For this reason, I will come back to this topic when I discuss the hybrid theories at STAGE III. Indeed, the legal example will serve as the blueprint of what I call the *burden of proof* strategy.

Here, I shall instead spend some time with the best example for an exclusively falsificationistic approach I can think of: Areas of discourse which allow *faultless disagreement*s.

6.7 Example II: Faultless Disagreement

The phenomenon of *faultless disagreement* arises in discussions about personal taste, humor, and (arguably) moral judgments. These disagreements, sometimes called "disputes of inclination," are exemplified in the following exchange:

Albert: "Spinach is tasty."

Bertha: "Spinach is not tasty."

The intuitions that guide people in those cases say that Albert and Bertha are, on the one hand, in a state of disagreement about the tastiness of spinach. On the other hand, though, it seems that neither is at *fault* here, in the sense that both are somehow entitled to their assertions. The task of the faultless disagreement theorist is to bring these two intuitions into harmony.[11]

Despite a great recent interest in these disputes,[12] all of the existent explanations run into difficulties when they try to do full justice to those intuitions.

6.7.1 General Strategies and Their Problems

Brushing in very broad strokes, there are currently three main strategies for explaining faultless disagreement on the table: Contextualism, realism, and relativism (Terminology is not completely consistent; I will mostly follow C. Wright's exposition in Wright (2006)).

First, there is the view that the propositions expressed in the dispute carry a hidden indexical parameter in their logical form. That is, an utterance of "Spinach is tasty" does not constitute an assertion of the simple proposition, (Spinach is tasty), but rather of the indexed proposition (Spinach is tasty according to S's taste), where S is the speaker.

---

[11] Or, if the thesis is that there is no such thing as a faultless disagreement, to explain why we have these mistaken intuitions. See Stojanovic (2007) for an attempt.

[12] See Kölbel (2009) and references therein.

Such *contextualism* can easily explain the faultlessness of disputes of inclination, but tends to have a problem accounting for genuine disagreement. After all, the disputants are talking about different things altogether. The case is analogous to examples in which other, uncontroversially indexical expressions are at play: You cannot object to my saying "I am German" by saying "No, I'm not." Similarly, if I spell out the suggested content by saying "Spinach is tasty according to my standards of taste," or simply "I find spinach tasty," you will not contradict me by saying "No, I do not!"

It seems that contextualism might have to concede that there is no real dispute at all. But, this is not only going against our initial intuition. It also makes dialectical moves we are apt to employ in disputes of inclination seem completely beside the point. Wright points out the following:

> This suggestion is open to the charge that it distorts the meaning of what we intend to say when we give voice to judgements of taste. There is, for example, a challenge involved in the question: 'If, as you say, rhubarb is delicious, how come nobody but you here likes it?', which goes missing if the proper construal of it mentions an explicit standard-relativity in the antecedent (Wright 2006, p. 38).

Compare to this genuine challenge the following facetious one: "If, as you say, you are German, how come no one else here thinks he is German?" If the contextualist is right, then pointing out that someone is alone in their taste judgement has no more suasive force than pointing out that there are no other Germans in the room. It seems the contextualist has to concede that "I can be right no matter how idiosyncratic my view."[13] But then, it seems hard to explain why such challenges seem to carry some weight.

In contrast to contextualism, *realism* posits that assertions about taste have exactly the logical form they seem to have and that there are objective facts that make such assertions true or false. There is the rather radical view that spinach either is intrinsically tasty, or it is not. If spinach is tasty, then it is tasty regardless of what *anyone* thinks about the matter. Wright calls such realism "rampant."

A more likely version is that there are objective facts about tastiness, but that these facts depend on people in some way. A realist account under this conception might take "tasty" to be short for "tasty for the majority of people" or "tasty for a group of culinary experts." That would be what Wright calls *response-dependent realism*.[14]

Realism, whether rampant or response dependent, has no problem with explaining why we perceive discussions about taste as disagreements. It is just the same kind of disagreement as in other disputes, say about the properties of middle-sized objects: One of the contestants has to be wrong about the relevant matters. But as with disputes about middle-sized objects, it is hard to see how such a dispute could be

---

[13] Wright (2006), p. 40.

[14] Note that the rampant realist shares a problem with the contextualist: She too has to concede that one can be right, no matter how idiosyncratic one's view is. Response-dependent realism would not have to allow for such a possibility; probably, the view would even entail that an idiosyncratic but right taste judgement is impossible, unless the relevant responses are those of a very small number of experts who happen to disagree with most other people.

called faultless. If my assertion is wrong, then surely I have made some kind of mistake. One option is to say that even though what I said was wrong, I made no *cognitive* mistake, in that I could not have done anything to find out the truth. But, what our intuitions speak of seems to be a more thorough kind of faultlessness.

The last of the main strategies is *relativism*. The idea of the relativist is that the propositions expressed in disputes of inclination are not true or false simpliciter, but only with respect to a given *perspective*. What is true-for-me must not be true-for-you.

It is only a recent development that analytic philosophers consider relativism as a potentially coherent position, and the phenomenon of faultless disagreement is a major incentive for developing the position.[15]

Over and above the question whether genuine relativism can be made sense of, it is not undisputed that relativism can better account for disagreement than contextualism.[16] If true-for-me can differ from true-for-you, is there really any sense in which we are defending *opposing* positions? Furthermore, it seems hard to see how the relativist can answer Wright's challenge and avoid to allow for correct but completely idiosyncratic taste judgements.

6.7.2 Unfalsifiability and Analetheism

Instead of going into the details of these positions, let us see whether Dummett's falsificationism can offer an attractive alternative. The proposal is this: In disputes about taste, there is nothing more to the correctness of an assertion than the impossibility to disprove it.

The idea seems to fit the disputes of taste quite well. Applying it to the spinach case, we get the result that neither Albert nor Bertha is at fault, because neither one made an assertion that the other or anyone else could falsify.

Indeed, quite a similar (albeit non-constructive) proposal has been made already. J. C. Beall[17] has suggested that areas in which faultless disagreements occur might be areas in which an analetheic approach might be fruitful. Indeed, apart from the context of semantic paradoxes in which analetheism was introduced, this is the only other motivation of the view that I'm aware of.

Recall from Sect. 4.12 that analetheists hold that an assertion is correct iff the asserted statement is not false. Further, they are prepared to accept truth value gaps, statements that are neither true nor false. These two contentions led them to endorse a three valued logic, in which not only the top but also the middle value (the gap) was treated as designated.

Beall applies these ideas to taste talk by claiming that all assertions about taste are neither true nor false:

---

15 See the contributions to Garcia-Carpintero and Kölbel (2008); other motivations for relativism include the analysis of future contingents and epistemic modals.

16 Ref. Stojanovic (2007) argues that the two approaches are merely notational variants.

17 Ref. Beall (2006).

> The idea, in a nutshell, is just this: some claims—for example, those involved in 'disputes of inclination'—are such that there's no fact of the matter, and so such claims are at least not false (Beall 2006, p. 67).

Defenders of this view will side with contextualism, rampant realism and (presumably) relativism in that they all have to bite the bullet when it comes to Wright's challenge: No matter how offbeat and abnormal my view about matters of taste, I can be right in holding it. After all, my assertion will be neither true nor false. Any attempt to argue against my view, for example by pointing out that no one shares my predilection, is misguided and betrays a considerable misunderstanding of my assertion.

I can muster some sympathy for Wright's intuition that this is not right. I believe that not all disagreements about taste need be faultless, and that there are taste judgements that are fundamentally misguided. I will call such a wrong taste judgment *perverted*.[18] I'll give you an example of a judgment that is ethically unobjectionable but that I take to be perverted:

> Bile is tasty

Bile,[19] as it happens, is not tasty. If you were to say that it is, we would not be experiencing a case of faultless disagreement, and that is because you would simply be wrong. I'm not prepared to agree that we cannot argue about taste in this case.

6.7.3 Agreeing to Differ

I'm sure that anyone intuits a difference between the spinach dispute and the one about bile. But, one might hold, the difference is only a matter of the statistic unlikelihood of meeting someone with a taste for bile. It might be hard to imagine the disagreement about bile actually taking place, the objection goes, but if it were taking place it would be as faultless as the spinach dispute. To suggest that really a line has been crossed in the case of bile, I'd like to draw your attention to the following: In cases of faultless disagreement, a truce is easily reached:

> A: "Spinach is tasty."
>
> B: "No, spinach is not tasty!"

---

[18] Of course, this is a highly problematic term. However, something strong is called for to convey the sense of something completely, utterly unacceptable (and something that is more catchy than "completely, utterly unacceptable"). One problem is that the term "perversion" carries a connotation of moral wrongness. Even though moral discourse has sometimes been suggested to allow for faultless disagreement, I intend to restrict myself to the taste examples here.
The slope is a slippery one, though. If I sincerely uttered "Human flesh is tasty", then many would agree that this is an instance of perversion. The question, however, is where the perversion lies. I think it is not in the taste judgment itself (after all, human flesh might taste just like chicken). Rather, knowing what human flesh tastes like already proves my depravity.

[19] We all can think of even clearer and more convincing examples; however, I prefer not to spell any of them out here.

A: "Well, I guess it's true what they say, you just can't argue about taste."

B: "Right."

Instead of persisting in pointing out the difference in opinion, the parties will quickly agree to differ and move on to other topics. In the cases where I argue no faultlessness is present because one party's taste judgment is perverted, a truce will be much harder to be come by:

A: "Bile is so tasty."

B: "Are you crazy? Bile is definitely not tasty!"

A: "Well, you can't argue about taste, can you?"

B: "What? No one thinks that bile is tasty!"

A: "Except for me."

B: "I'm sorry, but you're just wrong there!"

If you think that your opponent's position is genuinely beyond the pale (i.e., perverted), you will have a hard time leaving it at an agreement to differ. Conversely, if there were no perverted taste judgments, you would never have reason to want to decline the truce.

Each of the few languages I know has pieces of folk wisdom to the effect of "you cannot argue about taste" that we might cite and then retreat from the futile discussion. However, in some cases, we just could not be content with that because we feel that the opponent has committed a graver mistake than just opposing our own personal taste judgment.

This gives an explanation of why a charge such as "If, as you say, X is tasty, how come that everyone else hates it?" can be a genuine challenge to one's view. If there is such a thing as perversion, the argument can be understood to insinuate that the reason no one likes X is that liking X is just perverted. This points to the obvious fact that a rampant realism about perversion is not an attractive view. More plausible is a kind of response-dependent realism; the relevant responses, though, are not to the question whether the people like X or not, but whether they find liking X perverted or not.[20] It is a response-dependent realism about the boundaries of perverted taste.

Given that the most plausible way to *falsify* an assertion of the form "X is tasty" is to establish that enough people not only disagree but believe it to be utterly unacceptable.[21]

6.7.4 Falsificationism About Taste

We can now contrast Beall's analetheism with what I would take to be a plausible adaptation of Dummett's falsificationism to the issue. Both views agree that the

[20] These judgements can and will change over time. What once had been taken by the majority to be unacceptable might have become acceptable today and vice versa.

[21] Clearly, to say that X is not tasty is not to claim that "X is tasty" is falsifiable, but that "X is not tasty" is not a perverted judgement.

correctness of an assertion about taste depends on the absence of some negative quality. In the one case, this quality is simply falsity, whereas in the other, it is a falsification. While such an absence is guaranteed on Beall's view, I would give substance to the notion of a falsification of a taste judgement by appealing to the notion of perversion. To falsify your opponent's claim is to find evidence that most people find his taste perverted.

Can this view accommodate the intuitions of faultlessness and disagreement that the others had trouble with? I would say that an exchange such as the one about spinach is faultless, because surely neither a taste nor a distaste for spinach is perverted. It would be a disagreement of sorts, but probably Albert and Bertha would soon agree to disagree. In that case, I believe there is not a lot of pressure on a view to render the exchange as a genuine disagreement.

The interesting case is when one party declines the truce. Here, it is easy to hold that a real disagreement is taking place. Can such a disagreement continue to be faultless, though? Surely, not if one of the parties indeed holds a perverted position. But what if neither one of the expressed views is perverted, but the contenders refuse to settle the matter amiably? Then, I would say that the disagreement is in one sense still faultless, namely in the sense that both were entitled to their original positions, but that a different kind of fault is displayed by the party (or parties) that declines an agreement to differ. The only situation in which such a refusal to settle the argument is correct is when the opposed position is perverted. If it is not, the one who declines to settle the argument by agreeing to differ is at fault.

Faultless disagreements, thus, are those disputes in which neither party holds an unfalsifiable view on taste and in which a truce has not been offered. However, once a truce is offered, the situation will either cease to be faultless (because declining the truce is to imply that the other's position is perverted), or it would not be a disagreement any more (because accepting the truce is to lay aside the disagreement).

This is, of course, only a rough sketch of how a falsificationistic response to the problem of faultless disagreement could go. My main aim, however, was not so much to solve this particular puzzle, but to suggest example areas of discourse in which Dummett's idea of correctness as unfalsifiability could be attractive. I think that areas that seem to allow for faultless disagreements are among the best stabs we can have at this.

6.8 Correctness and Truth

We have seen two examples in which it may be claimed that an assertion's content is determined by its falsification conditions. Let us now come back to the question of how exactly we want to characterize the *correctness* and the *truth* of an assertion. I'll again start with correctness and subsequently turn to truth.

Dummett proposed two alternative ways of characterizing the correctness of an assertion (see Sect. 6.5). The first was to say that an assertion is correct iff it is not falsifiable at the present moment, that is, iff none of the actual states of affairs are ruled out by it. This would lead us to a notion of correctness that is *pro tempore*, as it might well happen that the asserted statement is falsified later on.

The second option was to reserve the label "correct" for those assertions for which we have conclusive evidence that they will never be falsified. This kind of correctness is not temporally constrained in the sense that a correct assertion could later become incorrect.

In criminal cases, the defendant makes correct assertions in the sense that all have to take his word unless he is proven wrong. This he achieves by asserting something that is, at the moment of the trial, unfalsifiable. There is no need to supply any evidence that would preclude future falsification, so the *pro tempore* notion of correctness seems to suffice here.

How about assertions about taste? What if the only thing we can say is that a taste judgement has not yet been falsified, but might become so later? Assume that Peter, a food chemist, concocts a new flavor in his laboratory. We hear Peter proclaim: "This is really tasty!"

Is what he says correct? It clearly is correct in the sense that he has, at the present moment, no reason to withdraw his assertion.

So, as far as we are concerned with correctness, understood as the absence of the obligation to retract a statement, *pro tempore* falsifiability seems the notion to go with. On the other hand, if we are trying to give an account of *truth* that has anything to do with our intuitions at all, *pro tempore* unfalsifiability would not seem too alluring.

Let us assume our defendant has made a claim that cannot, at the time of the trial, be falsified. Is that enough to say that he speaks the truth? This seems to be a far cry from our intuitive notion of truth. But if we do not want to say that what he says is true, in what way are we taking his word?

But, what if it becomes clear that there will *never* be any other evidence beside his testimony? Would that lessen our hesitation to say that he speaks the truth? It seems not. So, if truth is our aim, there is not much gain in plausibility in going from *pro tempore* to recognizably eternal unfalsifiability.

And what about truth in matters of taste? The minimal requirement that is often placed on truth is that the disquotational schema should be satisfied for all statements A[22]:

(DS) "A" is true iff A.

But is it *true* that spinach is tasty? Am I really committed to agree to "'Spinach is tasty' is true" if I assert that spinach is tasty? To my mind, it is not at all clear that the disquotational schema should be applied to statements about taste. Indeed, I would make a stronger claim: Unlike in the legal example, talk about truth in such matters seems quite beside the point.

---

[22] Dummett at one point derides this assumption as a "Pavlovian reaction" (Auxier and Hahn 2007, p. 180) of philosophers talking about truth. Not that he himself had never displayed this reaction before.

It seems that we are at a point where the tight bond between truth and correct assertibility threatens to unravel. While correct assertibility seems to accord reasonably well to *pro tempore* unfalsifiability, truth seems, if it can or should be captured at all, to be closer to recognizably eternal unfalsifiability. As I said earlier, where the two notions come apart, I shall stick with correct assertibility as central notion in my investigations. Where truth should end up is an interesting question that I will address as well, but it will not my main concern. The examples so far, especially the example of taste talk, suggest that the question of correct assertibility might have a chance of being answered where the question of truth just leaves us perplexed.

There is another phenomenon that I have not mentioned before that displays a similar pattern. The indicative conditional makes it notorious hard for us to say when it is true, but quite easy to tell when it has to be retracted.

6.9 The Uncertain Truth Conditions of Conditionals

Suppose I said yesterday:

> If it rains tomorrow, I will stay at home.

Suppose further that the antecedent turns out to be false. It is nice and sunny, and I'm out all day.

What should we say about the truth of this conditional? Was it true? Was it neither true nor false? Did I maybe not even make an assertion, because utterances of such conditionals should only count as assertions when the antecedent is fulfilled?

Dummett's answer is that it simply does not matter how we respond to these questions (TOE, p. 11). No matter which choice we make, it would not alter the use of the statement. The only important thing to appreciate about my statement is that I meant to exclude the possibility that it would rain and that I left the house. Everyone who understands the statement understands that this is the possibility I mean to rule out. As this possibility has not come to pass, it is clear that I did not make an incorrect assertion and that I will not have to retract my assertion. To understand all, this does not presuppose any opinion on whether the statement was true or not.[23]

A similar line of thought has led classical logicians to accept the material conditional as a gloss for indicative conditionals. Of the four possible combinations in a classical truth table, only one option is clearly false: When the antecedent is true and the consequent false. The other three slots are marked "true," in accordance with their assumption of bivalence. Such a simple solution of course is unavailable for constructivists who reject that assumption. But for them, as for the realists, the content of an indicative conditional is circumscribed by the states of affairs in which it is falsified.

---

[23] WTM, p. 82.

6.10 Falsificationistic Systems of Logic

Conditionals thus are pushing us in the direction of a falsificationism of some sort. Note the difference between this effect and the effect the considerations about negation were having. Dummett argues that both conditionals and negations motivate the introduction of falsifications. However, the use we put them to is a different one. We need falsifications to be able to say what the verification conditions of negations are, so that we may give a *verificationistic* theory of meaning that includes an account of complex statements. On the other hand, with conditionals, we are not even *trying* to give verification conditions any more. Conditionals, so Dummett, make us appreciate that we only need to grasp the falsification conditions of statements in order to understand them. They are therefore an incentive, not only to introduce falsifications into the ingredient sense, but rather to revert to a full-blown, pure falsificationism. We shall see how well he fares which such a purely falsificationistic setup in the first chapter of Part III. It will be necessary to give the meaning of all the connectives in terms of falsifications and also to set up a notion of logical consequence.

The job of logic, construed in terms of recognizable incorrectness, can be seen to amount to this: Suppose that we have asserted a number of statements. Suppose further that we do not have to retract them because they cannot be shown to be incorrect. We would like to take these statements as premises, and we want to know which consequences we may assert without running any further risk of being shown incorrect. In other words, we would like to say that logical consequence transmits the *absence of incorrectness*. One should at this point note the similarity to the idea of non-falsity preservation that we saw in the exposition of both Belnap's and the analetheist's views in Chap. 4. Of course, they had no constructivist constraint on their notion of non-falsity.

In the present, constructive setting, this same idea gives us the following: In those cases where Dummett is right and the recognizably incorrect assertions are only those that are actually falsifiable,[24] the consequence relation transmits *pro tempore* unfalsifiability. As we will see soon, this indeed seems to be the thought that was guiding Dummett when he went on to build a logical system to go with his falsificationism.

6.11 Chapter Summary

This chapter started with an investigation of Dummett's claim that the incorrectness of an assertion is a notion that is prior to the notion of correctness. This, he told us, is because only incorrect assertions have a direct and systematic consequence for our linguistic behavior: When the incorrectness of an assertion is pointed out, then it must be retracted.

[24] As opposed to cases where the (possibly temporary) absence of proof suffices for recognizable incorrectness, such as mathematical statements.

Assuming this much (at least for the sake of argument), we were lead to the following question: How can we get to a notion of correctness or truth from here? We saw that there are actually two distinct options hidden in the story Dummett tells: We might say that an assertion is correct iff it can not be falsified, or that it is correct iff we can *show* that it can *never* be falsified.

The first problem in our attempt to decide between these options was this: *Neither* account seems very convincing in all generality. For example, a mathematical assertion has to be retracted if it is pointed out that the speaker does not know of a proof of it, or at least of how to obtain such a proof. Here, incorrectness seems to be more than just falsifiability, and correctness less than mere unfalsifiability, in which ever sense we construe that notion.

To give some examples that were better suited to Dummett's proposal, I first discussed the assertions a defendant makes at a criminal trial. Given that the burden of proof lies on the prosecution, the defendant will only have to retract those assertions that can be falsified. Thus, there is a real sense in which his assertions have to be taken to be correct as long as they are unfalsifiable.

The second example we saw was taste talk. The proposal unfolded to be this: You may say what you will about taste, as long as your assertion cannot be falsified. This I gave some substance by suggesting what it is to falsify such an assertion. A taste judgement will be falsified if it can be shown that the majority of people not only disagree, but take the expressed sentiment to be outright perverted.

Even after the inspection of these examples, however, any clear-cut definition of correctness and truth seemed hard to pin down. In the end, we saw Dummett arguing that it does not matter much that we have no clear grasp of what notion of correctness or truth is the right one. As long as we know in which cases an assertion has to be taken back, we understand it sufficiently. This argument gains important support from the indicative conditional. With assertions of such conditionals, it is chronically hard to pin down the situations in which they are true or correct, and in some cases, it is even unclear whether or not we should say that they are incorrect/false or not. However, we do know in which cases such an assertion has to be withdrawn.

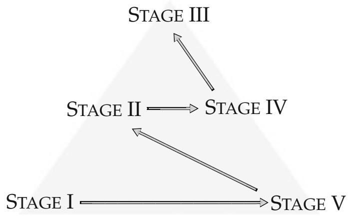
The conditional thus joins the negation as one of the driving forces in the campaign for falsifications. However, the conditional is more than an incentive to allow falsifications in the ingredient sense. While negation would rest content with having pushed us up to STAGE II, the conditional seems to call for a wholesale move to the right side of the pyramid, into the realm of falsificationism.

It will not come as a surprise that negation and the conditional will receive much attention in the third part, when I turn to the question which logical systems are in harmony with the ideas of the different stages. A general question about those logics was answered in the last section of this chapter: Which is the property that we should have the falsificationistic consequence relations transmit? The answer was that they should transmit *pro tempore* unfalsifiability. This is the positive constructive notion corresponding to the only property that has a direct influence on retraction: The recognizable incorrectness of an assertion.

Part III
Logics

Introduction to Part Three

In this third and last part, you will finally get to see some concrete proposals on the logical side. I will go through the Stages of the Pyramid, one chapter per stage. To set up the itinerary, I repeat the pyramidic diagram, this time with arrows indicating the order of the following chapters. It will make sense not to go through the stages numerically, but to zig-zag our way up to the apex. The chapter on Stage III theories will draw on many of the ideas and technical material that will come up at the lower levels, and it makes a lot of sense to treat the dual positions (I and V, II and IV) back to back.



In fact, I will not spend a chapter on Stage I. Not that there is nothing more to say about a purely verificationistic theory of meaning. Far from it, the list of topics that would need more attention is enormously long, even if we only consider logical matters. However, I take it that intuitionistic logic responds as well as any alternative to the needs of this stage, and I am content to leave it at that for the purposes of this book.

Chapter 7, therefore, deals with Stage V. This is the pure falsificationism that Dummett proposed in WTM, and I will introduce you to the logic he suggested there. It seems he thought it up on the spot when he wrote the essay, but elsewhere the logic is known as *dual intuitionistic logic*. It is, of course, dual to intuitionistic logic. The only question is: Dual in what sense? I will answer that question and point out a feature of dual intuitionistic logic that Dummett missed: It is paraconsistent, and that, I'll argue, is a good thing. In the end, however, I will claim that a purely falsificationistic account of logical constants is not much more satisfying than a purely verificationistic one. Thus, it will be time to move up to the level of the expanded theories.

Swinging back to the left side, I will spend Chap. 8 on Stage II, the expanded version of verificationism. The logic I will suggest at this stage will be one of the Nelson logics, namely the one called N_3. This is a constructive logic that allows for genuine gaps between verifiability and falsifiability, but not for gluts (the glutty variant is called N_4). It employs the toggle negation I introduced in Chap. 5. This negation switches back and forth between verification and falsification, and therefore validates all double negation laws.

Regarding the other connectives, the conditional will require the most amount of comments. I'll play through some possible variations, but in the end I will stick with the classic Nelson conditional. What might be most remarkable about this conditional is that it does not validate contraposition and *modus tollens*. These inferences are normally taken to be quite essential to a good conditional, so we will have to find out whether these failures can be made plausible or not.

After we are satisfied with the expanded verificationism of Stage II, we will in Chap. 9 turn to the equally expanded falsificationism at Stage IV. As for the models and the definitions of the connectives, we will find that we can use the very same Nelson models from Stage II. The only difference will be the way in which logical consequence is defined. Whereas at Stage II, verifiability was transmitted, here we will transmit nonfalsifiability, just as we did at Stage V.

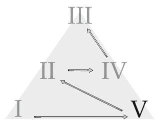
I will then turn to a problem that was already making an appearance at Stage V: falsificationism lets speakers make assertions that we wouldn't take to be correct, even if correctness is interpreted as weakly as we are doing anyhow on the right side of the pyramid. Examples include outright self-contradictions. Even if all we ask is that an assertion should not be falsified, is an assertion such as "I got into the car, and I did not get into the car" really correct because there is no evidence that constructively falsifies it? I suggest a notion of incoherence that is different from falsifiability, but still makes an assertion objectionable. I then try to capture that notion in the formal system we've been working with.

Finally, in Chap. 10 we arrive at the top, Stage III. Here, I will go through the different forms that a thorough blending of verifications and falsifications can result in. On the logical side, I will introduce a mechanism to combine the systems of Stage II and Stage IV.

Likewise, there will be little new to be said about the discourse separation strategy. To get a better idea of how the *burden of proof* strategy is supposed to work, I'll talk in more depth than I have until now about the example of legal discourse, which is both an application of and the template for a more comprehensive version of the strategy. I will also revisit the CV&IF strategy and probe its plausibility. To wrap it all up, I will get back to the question of how truth and falsity might fit into the various pictures I have sketched.

Chapter 7
Stage Five: Pure Falsificationism and Dual Intuitionistic Logic

7.1 Chapter Overview



Now, finally we are ready to inspect the first non-intuitionistic system that the introduction of falsifications brings us. As mentioned several times before, I start at the lower right hand corner of the pyramid, at Stage V. This is the purely falsificationistic theory, in which we have only falsifications to work with.

One reason for beginning with this stage is that the logic we will be looking at here is in an important sense the most similar one to intuitionistic logic. This similarity stems from the fact that it is, in a way that we will come to see presently, dual to intuitionistic logic. For that reason, it is known in the literature as *dual intuitionistic logic*.

A further reason to start here is that this logic is the only one in Part Three that was actually considered by Dummett. The systems of the later chapters are what I offer by way of rational reconstruction of his arguments. By watching closely how Dummett sets up the logic for this stage, we can learn much that will be useful in those later chapters.

I will, however, soon move away from the particular semantic guise in which he presents the logic. I believe much clarity can be gained by translating his proposal into a Kripke semantics that is very similar to the semantics for intuitionistic logic which I gave in Sect. 3.7. As we saw there, the Kripke semantics can only aspire to do more for the theory of meaning than a purely algebraic semantics if we supply it with a more intuitive interpretation. In the intuitionistic case, we used the BHK interpretation, which gave the meaning (that is, the proof conditions) of complex statements in terms of proof conditions of the constituent statements. Here, I will try to give you a similar BHK-style interpretation. Only this time, its task will be to give the *falsification* conditions of complex statements in terms of the *falsification* conditions of the constituents.

To give such a definition for conjunctions and disjunctions will be quite straightforward. It will get interesting when we get to the conditional. Dummett defines the

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_7,

conditional materially.[1] I will argue that this approach gives us nothing to be content with, and I will try out some alternatives. In the end, however, nothing completely convincing will turn up here.

Next, we will get to negation, which is the point at which the duality to intuitionistic logic comes out clearest. The characteristic failures of intuitionistic logic will actually hold, but dual ones will fail. So, for example, we will find double negation elimination (DNE) to be valid, but double negation *introduction* (DNI) to fail.

Among the peculiarities of the logic is one that Dummett missed: Explosion is among the inferences that are not valid any more. That is, the logic is a paraconsistent logic. On the other hand, the LEM will hold. I will argue that this is a very satisfying result and that we should expect both of these features of any logic for a falsificationistic theory.

However, giving the conditions for the BHK-style interpretation will not be trivial in this case. Indeed, the problems will be very similar to the problems we have seen arising with intuitionistic negation in Sect. 5.3.

7.2 Dummett's Falsificationistic Logic

Dummett's characterization of the logic that is to accompany his falsificationism is rather condensed; in fact, it is quite feasible to quote here (almost) everything he has to say about the formal features of the logic.

Here is Dummett's analysis of the logical consequences of a move to falsificationism[2]:

> Let us write f_A for the set of recognizable states of affairs in which A is falsified, and $f^\perp$ for the set of recognizable states of affairs which preclude the occurrence of any state of affairs in f. Plainly, $f \cap f^\perp = \emptyset$, $f \subseteq f^{\perp\perp}$, and, if $f \subseteq g$, then $g^\perp \subseteq f^\perp$; hence $f^\perp = f^{\perp\perp\perp}$. We may also assume that $(f \cup g)^\perp = f^\perp \cap g^\perp$ and that $f^\perp \cup g^\perp \subseteq (f \cap g)^\perp$. It seems reasonable to take $f_{\multimap A} = f_A^\perp$, $f_{A \vee B} = f_A \cap f_B$, $f_{A \wedge B} = f_A \cup f_B$ and $f_{A \to B} = f_B \cap f_A^\perp$, and to define $A_1, \ldots, A_n \vdash B$ as holding when $f_B \subseteq f_{A_1} \cup \ldots \cup f_{A_n}$, so that $\vdash B$ holds just in case $f_B = 0$; on this definition, however, $\vdash A \to B$ may hold when $A \vdash B$ does not. On this basis, we have $\multimap\multimap A \vdash A$ and $\vdash \multimap\multimap A \to A$, but not $A \vdash \multimap\multimap A$. We also have $\vdash A \vee \multimap A$, $\vdash \multimap (A \wedge \multimap A)$, $\multimap (A \wedge B) \dashv\vdash \multimap A \vee \multimap B$ and $\multimap (A \vee B) \vdash \multimap A \wedge \neg \multimap B$, but not $\multimap A \wedge \multimap B \vdash \multimap (A \vee B)$. However, I do not feel at all sure that this approach is correct (WTM, p. 83).

Without further comment, these claims are bound to puzzle the unassuming reader. In the following sections, I will unpack this to make it more perspicuous what is going

---

[1] By "material conditional," I mean a conditional $A \to B$ that is equivalent to $\neg A \vee B$. It will turn out that the negation in dual intuitionistic logic is an intensional one that looks to other worlds. Therefore, this material conditional will likewise not be determined by the truth value of A and B at the present world alone, which might be another way in which "material" could be understood.

[2] In the following quote, I write $\multimap$ where Dummett uses $\neg$. I will use the former sign throughout the thesis to denote the negation that is defined only in terms of falsifications. I also adjust Dummett's notation following I. Rumfitt's suggestion in Rumfitt (2007) and write $f^\perp$ instead of f'. Quite generally, Rumfitt's article is the most extensive discussion of Dummett's falsificationism.

on. To give it away already, I will end up sharing the doubt Dummett expresses in the last sentence of the quote.

7.3 Falsificationistic Semantics

Unlike the BHK semantics for intuitionistic logic, in the above quote the meanings of statements are not specified directly in terms of constructions of some sort (proofs, verifications, or falsifications). Rather, Dummett talks about *recognizable states of affairs* in which statements are falsified. In Rumfitt (2007), Rumfitt has made the sensible suggestion to interpret these as *states of information* or *stages of investigation,* such as we find them represented by the worlds in Kripke semantics.

Following up on this analogy, Rumfitt goes on to reinterpret Dummett's semantics in a way that is closer to the Kripke semantics we have seen in Chap. 3. Below I will present a semantics, that is, though superficially different, equivalent to Rumfitt's reconceptualization in all respects that are important here. We will be mainly looking to this Kripke semantics to get a better understanding of what this logic is about and especially of what we are to make of the connectives.

However, before we do so, let us stick with Dummett's original presentation at least long enough to be sure what he wants to preserve in his logic. I claimed at the end of Chap. 6 that the property that is to be transmitted from premises to conclusion should be *pro tempore* unfalsifiability. In the quote, we see Dummett defining "$A_1, \ldots, A_n \vdash B$ as holding when $f_B \subseteq f_{A_1} \cup \ldots \cup f_{A_n}$." f_B is the set of states of affairs (or states of information) in which B is falsified. This set is required to be a subset of $f_{A_1} \cup \ldots \cup f_{A_n}$. That is the set of states of affairs in which at least one of the premises is falsified. This tells us that $A_1, \ldots, A_n \vdash B$ holds if whenever all the premises are unfalsified in the current state, so is the conclusion. This is just what I meant by saying that the consequence relation transmits *pro tempore* unfalsifiability.

7.3.1 Rumfitt's "Safe Assertibility"

The alternative to *pro tempore* unfalsifiability as central notion we saw in the last chapter was recognizably eternal unfalsifiability. In discussing Dummett's falsificationism, Rumfitt has called recognizably eternally unfalsifiable statements *safely assertible.* The reason for that is obviously that a speaker who asserts such a statement can be secure in her knowledge that she will never have to retract it. Rumfitt believes that this notion is also the right one to tie *both* truth and correct assertibility to. Therefore, for him logical consequence should transmit recognizably eternal unfalsifiability, as well.[3]

Except for the definition of logical consequence, Rumfitt accepts all the definitions in the footnote quoted above, with the addition that he takes $f_A^{\perp}$ to denote the states

[3] However, he also writes that "the condition for some premises to entail a conclusion should be that a speaker who asserts the conclusion incurs no additional exposure or liability to future criticism or

of affairs in which statement A is safely assertible, because these are the states in which something precludes the falsification of A.

He further suggests we define $A_1, \ldots, A_n \vdash B$ to hold iff $f^{\perp}_{A_1} \cap \ldots \cap f^{\perp}_{A_n} \subseteq f^{\perp}_{B}$. Those states of affairs in which we can recognize that something precludes the falsification of each premise are a subset of those that preclude the falsification of the conclusion, i.e., whenever all of the premises are recognizably eternally unfalsifiable, so is the conclusion. Interestingly, Rumfitt shows that this redefinition of logical consequence will yield classical logic. He concludes that this offers an "anti-realism without the tears of an unfamiliar logic".[4] What is more, Dummett, in his reply to Rumfitt's essay, accepts the redefinition of logical consequence and writes that "safe assertibility seems to me the best notion of truth that can be attained within a falsificationist theory of meaning" (Dummett 2007, p. 697).

In the last chapter, I concluded that while recognizably eternal unfalsifiability might indeed be a better candidate for truth, it does not seem to be the better alternative when it comes to assertibility. Therefore, I will keep with Dummett's original proposal for now. I will come back to Rumfitt's idea in Sect. 9.4.

7.4 Kripke Semantics for Dual Intuitionistic Logic

Having established that we will concentrate on the transmission of *pro tempore* unfalsifiability, I will now move away from the semantic guise that Dummett chose and toward the Kripke semantics. The logic that he characterizes is, as far as I know, not found elsewhere in the literature. However, it is so closely related to what is known as *dual intuitionistic logic* that I have decided to simply call it by that name.

The difference lies in the conditional: Dummett in effect defines a material conditional. If you take a look at the clauses in the quote, you will realize that $A \rightarrow B$ and $\neg A \vee B$ are equivalent. Dual intuitionistic logic,[5] on the other hand, is normally presented without a conditional, but with a somewhat strange connective called "co-implication", a binary connective which I will render as $\prec$. A third version is what Priest calls *da Costa logic*,[6] which adds an intuitionistic conditional (of sorts) instead of the material one. What all these logics share is their conjunction,

(Footnote 3 continued)

remonstration that he has not already incurred by asserting the premises." (Rumfitt 2007, p. 669). In the present context, to me this actually sounds exactly like the transmission of *pro tempore* unfalsifiability, though the rest of the essay makes it clear that this is not what Rumfitt has in mind.

[4] Rumfitt (2007), p. 687.

[5] See Urbas (1996), Shramko (2005) and Miller (2006) for presentations. Miller also calls the logic *Brouwerian logic,* as the logic is related to Brouwerian algebras. However, neither logic nor algebra seem to have all too much to do with Brouwer's work.

[6] See Priest (2009). The title is to honor N. da Costa, one of the pioneers of paraconsistent logic. As in the case of Brouwer, however, the system has relatively little to do with da Costa's actual research. If one were to name the logic after its inventor, it might have to be called *Popperian logic*.

disjunction, and negation fragment. I will write $\mathrm{DIL}_{\rightarrow}$, $\mathrm{DIL}_{\prec}$ and $\mathrm{DIL}_{\supset}$ to keep them apart, and if the conditional is not at issue I will write simply DIL to mean the conjunction, disjunction, and negation fragment with any or none of these conditional-like connectives.[7] I will later discuss yet another conditional, $\supset_{\mathrm{Tol}}$, which is not found in the literature.

As the name *dual intuitionistic logic* as well as the examples of valid and invalid inferences in Dummett's quote will already have suggested, we are dealing with a dual to intuitionistic logic, in some sense of "dual." I have already remarked that "the dual of X" is not always as unambiguous as some writers will make you think. However, *some* form of dualization will take you to dual intuitionistic logic, almost regardless of which semantic or syntactic characterization of intuitionistic logic you choose.[8] Priest (2009) gives a quite comprehensive overview over these different dualization processes. I will first show how dualization works for the Kripke semantics and then interpret the result in detail in the later sections of this chapter.

7.4.1 Kripke Semantics: The Basic Setup

As in intuitionistic logic, we will start out with a partially ordered set of worlds, and we will take these worlds to be information states. The inquiring subject runs through these stages as she is engaged in her inquiry. So, a dual intuitionistic model

---

(Footnote 6 continued)

At least Popper (1948) contains the earliest description of it that I know of. However, the *idea* to dualize intuitionistic logic is quite a bit older; in fact, in a certain sense it is older than intuitionistic logic itself. The prize question that asked for the axiomatization of intuitionistic logic already asked, as a second part of the task, to dualize the resulting system:

"(...)It is asked to

1. construct (a logical system that captures Brouwer's ideas as far as possible)
2. to investigate whether from the system to be constructed a dual system may be obtained by (formally) interchanging the *principium tertii exclusi* and the *principium contradictionis*. " (See Troelstra (1990), p. 4)

It is not known why the author of the question, G. Mannoury, might have been interested in such a dual system, nor whether Heyting or anyone else ever seriously tried to answer the second part of the question. Heyting's original entry is lost, and the resulting publications make no mention of a dual system. Popper and the later developers of dual intuitionistic logic seem unaware of Mannoury's challenge.

[7] In principle, it is possible to use all three connectives at a time, which would give us $\mathrm{DIL}_{\rightarrow,\prec,\supset}$, but I will not have occasion to do so. In effect, this would amount to Heyting–Brouwer logic with a primitive material conditional and two definable negations. I will say more about Heyting–Brouwer logic below on p. 142.

[8] Indeed, Dummett's quote can be read as a topological characterization, where states of affairs are modeled as closed sets (in intuitionistic logic, they would be open sets, which marks the duality in this case).

will again be a structure, $[W, \leq, v]$, W a non-empty set of information states, and $\leq$ a partial order. Intuitively, $s \leq t$ will again mean that s is a state prior to t, or that t is a possible development from s. Also, v is again a valuation function that assigns exactly one of the values 0 and 1 to all statements at all worlds.

This time, however, the job of the inquiring subject is to falsify more and more statements, rather than to prove or verify them. This will mean that, as her insights accrue, more and more statements turn their value from 1 to 0. The value 0 here will mean that the statement is falsified,[9] the value 1 nothing more than that it is not yet falsified. In particular, 1 does not mean that the statement in question is verifiable, and it does not mean that it is true for now and evermore. A statement that has value 1 at a certain point can become falsified in any of the later stages. Consequently, we will have to give up the requirement that value 1 projects into the future.

However, if something is truly falsified, we may presume that it stays falsified. This is one of the points in which we make what is maybe an oversimplification, but it is one that is exactly as grave as the one that the intuitionist makes once she leaves the field of mathematics and advocates her logic for empirical research. Here too, one might wonder whether a verification will invariably hold up for ever more. The question is whether verifications and falsifications are conclusive, and also whether we want to count a statement verified (or falsified) when it had been verified (or falsified) at an earlier time which no one remembers. Let us assume for the rest of this chapter that the answer to this question is yes on both counts.[10] I will come back to this point in the next chapter.

This leaves us with the following idea: It is the value 0, not 1, that we should expect to project into the future. Let us make this a formal requirement:

For each p: if $w \leq w'$ and $v_w(p) = 0$ then $v'_w(p) = 0$.

The exchange of 0 and 1 in the heredity requirement is the first step in dualizing the semantics. In effect, it amounts to an inversion of the order of the worlds in a model for intuitionistic logic. Observe that our valuation functions assign nothing but 1s and 0s to statements. That means that a statement that gets value 0 in an intuitionistic model can never have had a different value at an earlier stage. The only other value it might have had would be 1, but then, due to the persistence of truth, it would still have value 1, not 0.

This means that we can keep the intuitionistic models and just flip the order around to get models that satisfy the new heredity requirement. In order to do this, we simply choose to keep reading $s \leq t$ as "s precedes t," but let $\leq$ denote from now on the

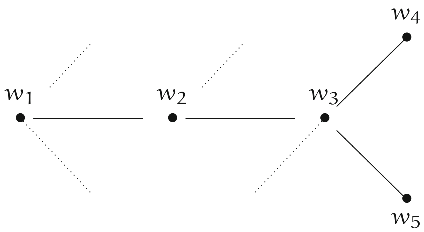
---

[9] To be more precise, I should say that the statement is falsifi*able*, in exactly the same sense as I explained the intuitionistic value 1 to mean that a mathematical statement is prov*able* in Sect. 3.8.4. I will allow myself the terminological sloppiness because "falsifiable" is potentially even more misleading than "falsified" if you do not keep the precise interpretation of the modality in mind.

[10] Rumfitt (2007), p. 663 claims that the steady accumulation of information is not only clearly an idealization, but actually goes against the laws of nature. Specifically, as the Second Law of Thermodynamics predicts a steady increase in overall entropy, the amount of available information (being structured in some way) cannot grow indefinitely. I shall not worry about this particular problem, neither here nor later.

converse of the intuitionistic accessibility relation (the converse of a partial order is still a partial order, so there is no problem in doing that).

Here is a diagram of a sample model:



	w_1	w_2	w_3	w_4	w_5
p	1	0	0	0	0
q	1	1	1	0	0
r	0	0	0	0	0
s	1	1	1	0	1

Just like the diagram of the intuitionistic model had the value 1 projecting into the future, this one preserves value 0. Note that if you read it from right to left instead of from left to right, you get a model in which, again, value 1 projects forward.

Before we turn to the clauses for the connectives, let us define logical consequence. Again, we want to transmit *pro tempore* unfalsifiability. That is, we want to transmit the *absence of value 0:*

$\Gamma \vDash A$ iff in every model and every $w \in W$, if $w \nVdash_0 B$ for every $B \in \Gamma$, then $w \nVdash_0 A$.

But since v is a total function, $w \nVdash_0 A$ and $w \Vdash_1 A$ are equivalent in this semantics. This means that we can reformulate the definition of the consequence relation so that it transmits value 1:

$\Gamma \vDash A$ iff in every model and every $w \in W$, if $w \Vdash_1 B$ for every $B \in \Gamma$, then $w \Vdash_1 A$.[11]

Note that this is exactly the same definition that we saw in our characterization of intuitionistic logic. The difference is that in intuitionistic logic, 1 stood for "provable," whereas here it stands for "not (yet) falsifiable."

[11] An early word of warning: The equivalence between $w \nVdash_0 A$ and $w \Vdash_1 A$ will break down from next chapter on, when we will start to consider *partial* valuation functions. That will of course mean that these two definitions of logical consequence would then lead to different logics.

7.4.2 The Connectives

We are right in the middle of our dualization process. Up to now, we have exchanged value 1 for value 0 in the heredity constraint, but we have kept the definition of logical consequence as it was in intuitionistic logic.[12] Now, we move on to the connectives. Here are the clauses we had seen in the chapter on intuitionistic logic again:

IL: For all $w \in W$,
$w \Vdash_1 A \wedge B$ iff $w \Vdash_1 A$ and $w \Vdash_1 B$
$w \Vdash_1 A \vee B$ iff $w \Vdash_1 A$ or $w \Vdash_1 B$
$w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \Vdash_0 A$ or $x \Vdash_1 B$
$w \Vdash_1 \sim A$ iff for all $x \geq w$, $x \Vdash_0 A$

And here are what are normally taken to be the clauses that are dual to these. You will see the conditional replaced by something called *co-implication* and denoted by the symbol $\prec$. I will get to the question of what it is supposed to mean later, for now the formal definition will have to do.

$\text{DIL}_{\prec}$: For all $w \in W$,
$w \Vdash_1 A \wedge B$ iff $w \Vdash_1 A$ and $w \Vdash_1 B$
$w \Vdash_1 A \vee B$ iff $w \Vdash_1 A$ or $w \Vdash_1 B$
$w \Vdash_1 A \prec B$ iff there is an $x \geq w$, $x \Vdash_1 A$ and $x \Vdash_0 B$
$w \Vdash_1 \multimap A$ iff there is an $x \geq w$, $x \Vdash_0 A$

Even though these clauses are often simply given as “the dual” to the intuitionistic ones, and even though it is clear that some things have been flipped and switched, it is actually quite an elaborate business to spell out what *exactly* has to be flipped over.

One way to give a recipe is this:

(i) Exchange 1 and 0 in the heredity constraint, but nowhere else
(ii) Leave the extensional connectives as they are For the intensional clauses,
(iii) switch universal and existential quantifiers in the metalanguage
(iv) exchange ‘and’ and ‘or’
(v) switch the values of antecedent and consequent of the dual of $\supset$.

Now, observe something that follows from the fact that $\nVdash_0$ and $\Vdash_1$ are equivalent: We can easily rewrite the clauses for the connectives as falsification conditions, i.e., conditions for a complex statement to receive value 0 at a world. This seems quite fitting for a falsificationistic theory. Here are the clauses:

[12] Note also that in some presentations, the heredity constraint is left as it is (i.e., projecting 1 forwards) and the relata of $\leq$ get switched around. In view of our interpretation, though, this would force us to read $x \leq y$ as “stage y precedes stage x,” which is not the most natural thing to do. In the end, it is a matter of taste, but the different ways of interpreting the relation are apt to cause some additional confusion.

DIL$_{\prec}$: For all $w \in W$,
$w \Vdash_0 A \wedge B$ iff $w \Vdash_0 A$ or $w \Vdash_0 B$
$w \Vdash_0 A \vee B$ iff $w \Vdash_0 A$ and $w \Vdash_0 B$
$w \Vdash_0 A \prec B$ iff for all $x \geq w$, $x \Vdash_0 A$ or $x \Vdash_1 B$
$w \Vdash_0 \multimap A$ iff for all $x \geq w$, $x \Vdash_1 A$

In fact, there is another way to dualize the intuitionistic semantics that will bring us directly to these clauses:

(i) Exchange 1 and 0 in the heredity constraint and also in the clauses for the connectives, except for the clause of the dual of $\supset$ (and also not in the definition of logical consequence)
(ii) exchange 'and' and 'or' in the clauses for the extensional connectives.

This looks a bit more concise than the above list, but still it seems clear that there are other ways to go about this dualizing business that would be more obvious but lead to different clauses. For example, we might wonder why we have to treat the dual of the conditional differently from the other connectives.

In fact, P. Schroeder–Heister [13] argues that we *should not* treat them differently and that the actual dual of $\supset$ is the converse of what we call co-implication (he also thinks that this terminology is flawed). I believe he has a point in this, even though his is clearly the minority position and it is probably too late to make the adjustment against the will of the critical mass.

Not that the adoption of $\prec$ as the dual of $\supset$ is accepted by the other authors without any rationale. We get the following pleasing symmetry:

In intuitionistic logic, we have that

$A \wedge B \vDash C$ iff $A \vDash B \supset C$,

and in DIL$_{\prec}$we find:

$A \vDash B \vee C$ iff $A \prec B \vDash C$

This makes $\prec$ attractive, even if the dualization of the Kripke clauses suffers a setback in perspicuity. All this is mainly said to substantiate my earlier claim: What "the dual of X" is often not pristinely clear.

Let us turn our attention back to the second set of semantic clauses, which defines explicitly when a complex statement is given the value 0. As I said, this way of presenting the clauses seems more in line with our aim: to explain the meaning of the connectives by explaining what should count as a falsification of the statements in which they occur. I shall try to read off an even more specific BHK-style interpretation in terms of falsifications in the next section.

However, already we can see that some progress has been made with respect to Dummett's original presentation. The clauses for conjunction and disjunction are quite obvious reformulations of what we read in his quote. However, the clause for negation is telling us something slightly more tangible than his description of $f_A^{\perp}$

[13] Schroeder-Heister (2009).

(i.e., the set of the states in which $\multimap A$ is falsified): "[T]he set of recognizable states of affairs which preclude the occurrence of any state of affairs in" f_A (i.e., the states in which A is falsified).

The worlds in which $\multimap A$ is falsified (receives value 0) are those in which there is no epistemic possibility of falsifying A. Therefore, in all accessible worlds, we find that A is assigned value 1. Now, this progress is not yet tremendous; we still lack a good idea of what it is that we might know that rules out the future worlds in which A is falsified. It will be the business of the BHK clause below to tell us, and to tell us *in terms of falsifications*. Still, we know a bit more about how to think of $f_A^{\perp}$ and the meaning of negations.

On the other hand, the clause we find for $\prec$ tells us nothing at all about how to model conditionals in DIL. I will return to the question what co-implication might mean, but this much is very clear from inspecting the clause alone: It neither is a contender for anything in the ballpark of a conditional, nor does it have anything to do with Dummett's proposal for the conditional.

He gave the following definition: $f_{A \to B} = f_B \cap f_A^{\perp}$. This, as I have already mentioned, is just the same as what we would get for $f_{\multimap A \vee B}$. Given what we have just learned about the representation of $f_A^{\perp}$ in our semantics, it is easy to give the clause for Dummett's material conditional. The clauses that correspond to his version of dual intuitionistic logic, $\mathrm{DIL}_{\to}$, are then:

$\mathrm{DIL}_{\to}$: For all $w \in W$,

$w \Vdash_0 A \wedge B$ iff $w \Vdash_0 A$ or $w \Vdash_0 B$

$w \Vdash_0 A \vee B$ iff $w \Vdash_0 A$ and $w \Vdash_0 B$

$w \Vdash_0 A \to B$ iff for all $x \geq w$, $x \Vdash_1 A$ and $w \Vdash_0 B$

$w \Vdash_0 \multimap A$ iff for all $x \geq w$, $x \Vdash_1 A$

This translates over into the positive notions (which are more useful for drawing inferences, as consequence is defined in terms of 1-preservation) as follows:

$\mathrm{DIL}_{\to}$: For all $w \in W$,

$w \Vdash_1 A \wedge B$ iff $w \Vdash_1 A$ and $w \Vdash_1 B$

$w \Vdash_1 A \vee B$ iff $w \Vdash_1 A$ or $w \Vdash_1 B$

$w \Vdash_1 A \to B$ iff there is an $x \geq w$, $x \Vdash_0 A$ or $w \Vdash_1 B$

$w \Vdash_1 \multimap A$ iff there is an $x \geq w$, $x \Vdash_0 A$

7.5 Some Characteristic Features of DIL

With Dummett's definition of the conditional, a large portion of the duality between intuitionistic logic and dual intuitionistic logic actually breaks down. However, if we concentrate on the conjunction-disjunction-negation fragments of these logics, we can see how the duality in the semantics affects the theorems and valid consequences in perfectly symmetric ways.

As we noted many times before, in intuitionistic logic, Double Negation Elimination fails to hold ($\sim\sim A \nvDash A$), as well as Excluded Middle ($\nvDash A \vee \sim A$) and one of the De Morgan Laws ($\sim (A \wedge B) \nvDash (\sim A \vee \sim B)$).

All of these inferences are valid in dual intuitionistic logic (swapping $\multimap$ for $\sim$). However, Double Negation *Introduction* ($A \nvDash \multimap\multimap A$) and the following De Morgan Law fail: $\multimap A \wedge \multimap B \nvDash \multimap (A \vee B)$. As Dummett noted in the footnote quoted above, the characteristic failures of intuitionistic logic seem to have dual failures in dual intuitionistic logic.

What about Excluded Middle, though? If we think about what might be the dual invalidity to $\nvDash (A \vee \sim A)$, we might as a first guess propose the invalidity of $\multimap (A \wedge \multimap A)$. However, just as Dummett says, this is indeed valid in dual intuitionistic logic as well.

I have discussed the question of the "real" dual principle to LEM in Sect. 4.11. The solution to the puzzle is that Explosion, $(A \wedge \multimap A) \vDash B$, is not a valid inference in dual intuitionistic logic (expressing the LEM as $B \vDash A \vee \neg A$ helps see the symmetry). In other words, dual intuitionistic logic is a paraconsistent logic.

The exact result of the duality in the semantics is the following duality in the sets of valid inferences of intuitionistic and dual intuitionistic logic: A single premise inference is valid in one logic iff the inference one gets from exchanging (a) premise and conclusion and (b) conjunctions and disjunctions is valid in the other logic.

	IL	DIL
$\neg\neg A \vDash A$	n	Y
$A \vDash \neg\neg A$	Y	n
$B \vDash A \vee \neg A$	n	Y
$A \wedge \neg A \vDash B$	Y	n
$\neg(A \wedge B) \vDash \neg A \vee \neg B$	n	Y
$\neg A \wedge \neg B \vDash \neg(A \vee B)$	Y	n

7.5.1 Paraconsistency

If you look at the quote at the beginning of this chapter, you will notice that Dummett makes no mention of the fact that the logic is paraconsistent. As he had worked out some of the peculiarities of the logic, this seems to be an interesting oversight. Even more so as the failure of Explosion seems to have a much more immediate impact on metaphysical questions than, say, the failure of a De Morgan's law might have. As we had seen in Sect. 4.10, the modern view is that the validity of Explosion indicates that the semantic law of non-contradiction is assumed.

However, as is clear from the preface to TOE, Dummett had not presaged this modern turn of events. He took the validity or invalidity of $\neg(A \wedge \neg A)$ to be the formal counterpart to the (semantic) law of non-contradiction. As he *did* note in the quote about his falsificationistic semantics, this is valid in dual intuitionistic logic, just as we have seen it valid in other important paraconsistent logics, such as Priest's Logic of Paradox (LP).

What, then, should we be saying about the relation between dual intuitionistic logic and the law of non-contradiction, given that Explosion fails? The law of non-contradiction tells us that no statement is both true and false (No Gluts) and that no statement could be true if its negation is (Negation Incompatibility). In the last chapter (Chap. 6), we have seen that the question what falsificationistic truth and falsity might be is not at all easily answered. At least, we had at the end of the chapter an idea of what we would choose as a designated semantic value: *Pro tempore* unfalsifiability.

Lacking a good grip on truth and falsity, we might interpret No Gluts to be saying that no statement receives both a designated and an undesignated value. In this sense, there seem to be no truth value gluts in our semantics. There are only two semantic values, and each statement has one or the other at every world, but never both. The No Gluts part of the law of non-contradiction holds good, at least in that particular sense.

However, if we take the law of non-contradiction to be also saying that no statement can receive a designated value if its negation does, then it is surely false in the semantics for dual intuitionistic logic.[14] Interpreted this way, Negation Incompatibility and with it the law of non-contradiction fails.

Apart from the theoretical question whether the law of non-contradiction holds or not, I would like to suggest that we can appreciate intuitively that paraconsistency is a desirable feature for a falsificationistic logic. The situation is, once again, parallel to what we saw in the discussion of intuitionism. There, it was easily seen why a constructivist/verificationist should not assume the validity of the Law of Excluded Middle. The reason was that we simply have no right to assume that every statement or its negation should be verifiable.

Similarly, we have no right to assume that every statement or its negation is *falsifiable*. But to falsify a conjunction, as a look at the semantic clause tells us, is to falsify one of the conjuncts. Therefore, we have no right to assume that $A \wedge \multimap A$ is always falsifiable, that is, that such a statement always receives an undesignated value. But without making such an assumption, it would be reckless to use a logic that satisfies Explosion.

But what about the validity of LEM in dual intuitionistic logic? Well, this only shows that $A \vee \multimap A$ is never falsifiable. What would be needed to falsify such a statement would be a falsification of A, and a falsification of $\multimap A$. We have not yet seen exactly what a falsification of $\multimap A$ is, but intuitively it is clear that these two falsifications should not be obtainable at the same time.

[14] Here is a counter model:

Let $w_1 \Vdash_1 p$ and $w_2 \Vdash_0 p$. Then, the clause for negation gives us that also $w_1 \Vdash_1 \multimap p$.

It is then very plausible to assume that any falsificationistic logic (whether pure or expanded) should satisfy LEM and be paraconsistent.

7.6 A BHK-Style Interpretation

In Sect. 3.7.3, we had seen that the Kripke semantics for intuitionistic logic had to be supplied with a more intuitive explanation, such as the BHK interpretation. Otherwise, it could not quite play the role that Dummett had set out for such a semantic theory. Likewise, to explain what exactly our dual intuitionistic models have to do with falsifications, we should hope to find a BHK-style account of the connectives that informs us about their falsification conditions.

7.6.1 Conjunction and Disjunction

The first two items on the list, conjunction and disjunction, come quite naturally. The two clauses

$w \Vdash_0 A \wedge B$ iff $w \Vdash_0 A$ or $w \Vdash_0 B$

$w \Vdash_0 A \vee B$ iff $w \Vdash_0 A$ and $w \Vdash_0 B$,

and also Dummett's $f_{A \vee B} = f_A \cap f_B$ and $f_{A \wedge B} = f_A \cup f_B$ naturally suggest the following:

- c is a falsification of $A \wedge B$ iff c is a pair $(i, c1)$ such that $i = 0$ and $c1$ is a falsification of A or $i = 1$ and $c1$ is a falsification of B
- c is a falsification of $A \vee B$ iff c is a pair $(c1, c2)$ such that $c1$ is a falsification of A and $c2$ is a falsification of B

Intuitively, this sounds right as well. If you want to falsify Jones's assertion "I was at home, and I was talking to my wife," you should either strive to show that Jones was not at home at the time in question, or you should try to show that he was not talking to his wife. However, it is not necessary to show that Jones was lying on both counts, i.e., you do not have to show that he was neither at home nor talking to his wife.

On the other hand, if what Jones said was "I was either at the department or in the bar," you have to show that Jones was *neither* at the department *nor* in the bar. Falsifying one of the two claims will not be enough.

7.6.2 The Conditional

Now we proceed to the more exciting connectives, and we will start with the conditional. Remember that the conditional was one of the chief reasons to turn to falsifications in the first place. It was Dummett's contention that we know very well what a falsification of a conditional amounts to, so we should be able to spell out the BHK-style explanation straight off the bat:

- c is a falsification of $A \supset B$ iff c is a pair $(c1, c2)$ such that $c1$ is a verification of A and $c2$ is a falsification of B

We falsify a conditional by verifying the antecedent and falsifying the consequent, it is that easy. What a conditional excludes is that the antecedent is the case and the consequent isn't, and this BHK clause expresses this idea splendidly.

The problem, of course, is that this explanation does not only speak of falsifications, but of falsifications and verifications together. That is, it is clearly committed to what I called an expanded falsificationism. We will be happily employing this clause once we ascend to the next level of the pyramid. For now, though, it is out of bounds.

As a substitute, Dummett gives us the material conditional, $f_{A \to B} = f_B \cap f_A^{\perp}$. That is, the states in which $A \to B$ is falsified are those in which B is falsified and something precludes the falsification of A. We will talk about how to explain "something precludes the falsification of A" in terms of falsifications when we get to the negation. For now, let us assume that it will be possible to do so and ask this: Is this, intuitively, delineating the states in which we would say that a conditional is falsified?

Let us suppose that I, in an attempt to show off my fortune telling abilities, say to you on new year's eve: "If a black cat crosses your path tomorrow, you will not catch a cold all through the year." You are not quite sure what to make of this prediction, but resolve to watch out for black cats the next day. However, as the day unfolds the whole business slips from your mind. Two weeks later you come down with a bad cold. This is the point where you remember what I had said. You cannot remember whether you saw a black cat or not, and there is surely no one else who followed you around and paid attention to this matter. It seems this is enough to preclude future falsification (or verification, for that matter) of the antecedent. The consequent, on the other hand, is clearly falsified. That is, our current state of affairs is in the set $\text{f}_{\text{No Cold}} \cap \text{f}^{\perp}_{\text{Black Cat}}$, and thus in $\text{f}_{\text{Black Cat} \to \text{No Cold}}$.

But would we really say that you have falsified my assertion? Would it be right of you to call me and demand that I withdraw it?

> You said I wouldn't get sick if a black cat crossed my way on the first of January, and now I have a fever!
>
> I'm sorry and surprised to hear it; did you see a black cat, then?
>
> I don't know, but I'm sure you can't prove that I didn't!

It is clear that this is not a convincing way to get me to take back what I said. Even if you replied instead

> I surely did!,

knowing full well that I cannot falsify your claim, I could simply reply

> Oh, but your cold shows that you couldn't have seen one. Probably you've mistaken a small dog for a cat.

As we assume that we are uttering our claims under the falsificationistic norm, both of our assertions will stand. What is missing to force me to retract my conditional is clearly *the verification of the antecedent*.

In the last chapter, I had already voiced my doubts about the applicability of the falsificationistic program to *all* areas of discourse. Now we see that, in general, the purely falsificationistic account of conditionals is unsatisfying. But what about those cases in which I saw a better chance for the falsificationistic analysis, i.e., legal and taste talk? Is Dummett's account of the conditional a better one in these areas?

In the legal case, there seems to be no improvement to the example above. Let us suppose that the defendant, Smith, is insinuated to have threatened a certain Peterson, who has subsequently been murdered. The only evidence available for the claim that Smith threatened Peterson is the testimony of their mutual acquaintance, Johnson. However, Johnson is not too sure he is remembering everything about the incident correctly, the threat might well have been a figment of his imagination. It surely isn't established yet that the threat really was uttered.

However, suppose it is equally clear that it will never be falsified that the threat was uttered, because the three were alone in a sound proof room, we somehow know that the victim, Peterson, had left no records of the incident etc..

Now, Smith, who likewise isn't completely sure about the exact course of events any more, says this:

> If I really did threaten Peterson, then I was drunk.

As the trial unfolds, new evidence comes in. Smith had, shortly after the incident, been tested for alcohol in a routine road control. The test established that Smith hadn't had a drink the whole day.

This seems to speak against his earlier conditional no more than your cold did against my prediction. Smith will not have to withdraw his earlier assertion, but he will now have to say that he in fact never threatened Peterson. Only when it is verified that he really did threaten Peterson will the conditional have been falsified. So, once again we are in need of verifications in the ingredient sense here.

In the legal case, it was clear from the beginning that verification conditions would have some role to play, because at least the prosecution was judged against a verificationistic norm of assertion. An area of discourse that might be able to do without verifications altogether, I claimed, was taste talk.

Let us then consider a large group of boy scouts trudging through the woods. They come across a large clearing. The ground is covered with mushrooms of various shapes, sizes, and colors.

Adam says this:

> If any one of those mushrooms is tasty, then it's got to be that big brown one.

Bert takes a bite off of that same mushroom, chews a bit and spits it out: "Yuck! This is disgusting, absolutely disgusting!!"

All of the boys, being of the inquisitive kind, taste the mushroom, and the verdict is unanimous: Anyone who were to take a liking to this taste must be perfectly mad. The mushroom is completely unfit for consumption. Let us assume that this is enough to establish that we are in $\mathfrak{f}_{\text{Brown Mushroom is Tasty}}$.

Before they can turn their attention to the other mushrooms, a wild animal chases them away from the clearing and tramples all of the mushrooms to a pulp. Surely, the claim that some of them were tasty cannot be falsified any more. We are in $\mathfrak{f}_{\text{Brown Mushroom is Tasty}} \cap \mathfrak{f}^{\perp}_{\text{Some Mushroom is Tasty}}$.

May the boys turn to Adam and ask him to retract his conditional? Clearly not. He will simply say: "Well, if the brown one was bad, then none of these other mushrooms were tasty!".

What all of these examples have in common is this: The speaker is committed to the negation of the antecedent once the consequent is falsified. But that is quite different from having to retract the assertion of the conditional. All in all, Dummett's falsificationistic analysis of the conditional is a failure.

Might there be other possibilities to get to a purely falsificationistic conditional? Maybe we can take some inspiration from the intuitionistic solution to this problem.

Intuitionistic conditionals

The intuitionistic conditional was explained in terms of a procedure that turns one verification into another verification. Maybe, the thought suggests itself, the falsificationistic account of conditionals should invoke procedures that turn one falsification into another falsification.

However, to take the conditional to be a procedure that turns a falsification of the antecedent into a falsification of the consequent is surely the wrong way to go about this. A cue from *modus tollens* will immediately give us a better idea: Have the conditional turn falsifications of the consequent into falsifications of the antecedent. Let us symbolize such a connective by $\supset_{\text{Tol}}$, for "tollens."

Implementing this thought into the Kripke semantics in the same way as we did with the intuitionistic conditional, we get the following clause, expressing the idea that at each future state, either B has not yet been falsified, or A has been falsified:

$$w \Vdash_1 A \supset_{\text{Tol}} B \text{ iff for all } x \geq w,\ x \nVdash_0 B \text{ or } x \Vdash_0 A$$

But since in these semantics, $x \nVdash_0 B$ is equivalent to $x \Vdash_1 B$, we can rewrite this clause as

$$w \Vdash_1 A \supset_{\text{Tol}} B \text{ iff for all } x \geq w,\ x \Vdash_0 A \text{ or } x \Vdash_1 B,$$

which is just the same clause as the intuitionistic one!

However, something in our story here went fundamentally wrong: The account, even though it is given in terms of falsifications, points out the cases in which $w \Vdash_1 A \supset_{\text{Tol}} B$ is true of a world in a model. What we should like to know, however, is when $w \Vdash_0 A \supset_{\text{Tol}} B$ holds. This becomes perfectly obvious when we spell out the BHK-style clause. It is perfectly clear how the right hand side of the biconditional will come out, but what are we defining? If we had verifications around, we could write this:

- c is a verification of $A \supset_{\text{Tol}} B$ iff c is a procedure that converts each falsification of B into a falsification of A

But we are not interested in finding out what a verification of "If A, then B" amounts to, because we are trying to get by without verifications! Value 1 in the present semantics simply does not stand for a construction of any sort, only for the *absence* of a construction. The only constructive notion we have at our disposal is falsification; simply exchanging verification for falsification would give us this clause for some connective $*$:

- c is a falsification of $A * B$ iff c is a procedure that converts each falsification of B into a falsification of A

But this clearly is not what we are looking for. Incidentally, it is the correct BHK clause for the *co-implication*, to which I will turn in the next section.

So, how might we express the falsification condition for a conditional, seen as essentially a method of turning a falsification of the consequent into a falsification of the antecedent? Moreover, how can we phrase this condition only in terms of falsifications and events that preclude future falsification?

The best idea would seem to be this: Let the conditional be falsified iff the consequent is falsified and something obtains that precludes the future falsification of the antecedent. But this would just be $\rightarrow$, the material conditional above! We have seen that this conditional does not correspond well to our intuitions about when a conditional is falsified. Moreover, modus tollens is not even valid in $\text{DIL}_{\rightarrow}$, which would seem to be the minimum requirement for a conditional that purports to embody the spirit of this inference.

Let us get back to $\supset_{\text{Tol}}$ once more. The fact that this conditional aims to capture a positive, and not a negative constructive notion is not the only, and probably not even the gravest problem this account has. If we add it to DIL and subject it to the heredity constraint, then the delicate ecosystem of our world semantics collapses and we end up with classical logic: First, observe that we will be able to define classical negation $\neg A$ as short for $A \supset_{\text{Tol}} \multimap (B \vee \multimap B)$.[15]

But then all our worlds will look the same, because it is impossible for a statement to be falsifiable at a world if it has not *always* been falsifiable. To see this, suppose that $u \leq w$, $u \Vdash_1 A$. Then, we will have $u \Vdash_0 \neg A$, because of the heredity constraint $w \Vdash_0 \neg A$ and therefore, as $\neg$ is classical negation, also $w \Vdash_1 A$. As all the worlds look the same, they are nothing but decorative baubles. What we end up with is simply classical logic with two equivalent classical negations, $\neg$ and $\multimap$.

So, $\supset_{\text{Tol}}$ is, we have to concede, a failure as well. What we have just seen, however, is not the only way we might try to add something like an intuitionistic conditional to DIL. A feasible way to do it is this: First, remember that dual intuitionistic models can be seen as intuitionistic models with the accessibility relation turned around. Now, the trick is to have the conditional "look" in the direction into which the value 1 is projected, just as the original intuitionistic conditional did. That is, it is looking backward down the dual intuitionistic accessibility relation:

---

[15] Cf. Miller (2006).

$w \Vdash_1 A \supset B$ iff for all $x \leq w$, $x \Vdash_0 A$ or $x \Vdash_1 B$

or, as we are interested in the falsification conditions,

$w \Vdash_0 A \supset B$ iff there is an $x \leq w$ such that $x \Vdash_1 A$ and $x \Vdash_0 B$

This, in contrast to the above, gives no trouble like the collapse of dual intuitionistic negation into classical negation, or some such. The addition of this conditional is a conservative extension of DIL, which we may simply call $\text{DIL}_{\supset}$.

In fact, there is room for even further extension here. It is possible to add intuitionistic negation as well as co-implication, where the former will look backward to the past and the latter forward into the future. This comprehensive logic is known as *Heyting–Brouwer* logic.[16] The criticism of $\text{DIL}_{\supset}$ I am about to put forward concerns the intuitionistic conditional and extends also to Heyting–Brouwer logic.

The problem with $\text{DIL}_{\supset}$ is simply this: It is wholly unclear why the conditional should care about the worlds that lie in the past. What $A \supset B$ seems to tell us here is that B surely has not been falsified before A, if at all. This does not latch onto natural language "if...then..." constructions very well, and it does not translate into any falsificationistic BHK clause I can fathom. The criticism of $\text{DIL}_{\supset}$ and its Kripke semantics, then, is reminiscent of what Dummett said about the topological semantics for intuitionistic logic. The semantics is a "merely algebraic" one and has no clear connection to actual recognitional aspects, and I do not know of a different semantics that might do any better.

All this leaves us, I should say, rather unsatisfied. The conditional was a major incentive to move to pure falsificationism, but now it turns out to be much more of a liability for the view. We will see whether we can do any better in the case of negation, but first I should say something about the mysterious new connective, the co-implication.

7.6.3 Co-implication

We already know quite a bit about co-implication. We know how to define it in the Kripke semantics, namely by giving either of the following clauses:

$w \Vdash_0 A \prec B$ iff for all $x \geq w$, $x \Vdash_0 A$ or $x \Vdash_1 B$, which is equivalent to

$w \Vdash_1 A \prec B$ iff there is an $x \geq w$, $x \Vdash_1 A$ and $x \Vdash_0 B$

We have also just now seen the BHK clause which corresponds to this:

- c is a falsification of $A \prec B$ iff c is a procedure that converts each falsification of B into a falsification of A

This is an asset of some consequence, and it puts co-implication in a better position than the intuitionistic conditional we saw in $\text{DIL}_{\supset}$. And it beats $\supset_{\text{Tol}}$ hands down, because it does not interfere with the other constants.

[16] See Rauszer (1977, 1980).

Lastly, we can say in its favor that we know that it has some right to consider itself the official dual to intuitionistic conditional (even though this is not undisputed, as we saw on p. 133).

The only thing that is unclear is this: What are we supposed to do with it? What might a natural language notion be that we could analyze with its help? Glosses that have been proposed for $A \prec B$ are "A without B," "A minus B",[17] "A excludes B"[18] or "A but not B".[19]

To be honest, these do not seem very good glosses to me. But even if they were spot on, they are not the kind of logical constants that urgently await modeling, as, in contrast, "If A, then B" surely is. But we have seen from the start that co-implication cannot aspire to be an analysis of indicative conditionals. If anything, it is somehow the opposite of an indicative conditional. Maybe we can model assertions of the form "It is not the case that, if A, then B".[20] Again, having an analysis for this construction but lacking one for the simple indicative conditional is no reason to rejoice.

Co-implication, then, is not doing any harm, but it is not too useful either. That we may have it in our falsificationistic logic does not make up for the lack of a good conditional.

7.6.4 Negation

Let us at last take a closer look at the negation. We know this much: $\multimap A$ will be counted as falsified at a world in a Kripke model iff all conceivable courses of further investigation will leave A unfalsified. However, just as we had observed when we looked at the Kripke semantics for intuitionistic logic, this alone is not telling us enough. What constructively plausible reason could be given for the fact that we should model a state of information as one in which all accessible states leave a given statement unfalsified?

Even though we saw no natural gloss for the co-implication, $\prec$, here we can put it to some use. We can define negation in terms of it and a constant $\top$ that is never falsified, just as we used the intuitionistic conditional and $\bot$ to capture intuitionistic negation.

We define $\multimap A$ as $\top \prec A$. It is not hard to see that the clauses in the Kripke model support this identification, and it immediately gives us a BHK-style notion of dual intuitionistic negation:

- c is a falsification of $\multimap A$ iff c is a procedure which transforms each falsification of A into a falsification of $\top$,

---

17 Both in Restall (1997).

18 Urbas (1996).

19 Goodman (1981).

20 Seeing this, one might try $\multimap (A \prec B)$ for an indicative conditional. Things are starting to get a bit complicated here, but one can check that the cat/murder/mushroom examples that ruled out the material conditional will likewise disqualify $\multimap (A \prec B)$.

where $\top$ is something that is clearly and beyond doubt true, or maybe some element of "the set of things that are fundamental to our inquiry," as Priest[21] puts it.

Of course, the notion of a "falsification of $\top$ " is then just as problematic as the idea of a "verification of $\bot$"; the story you are about to read is a faithful, if dual, remake of what you read before. I will keep it short.

As you would expect if you recall Sects. 3.6 and 5.3 , we need a $\top$ that satisfies two conditions: A falsification of a true statement should lead to a falsification of $\top$, and a falsification of $\top$ should lead to a falsification of any statement whatsoever.

If we only found a $\top$ that satisfied the first of these requests, we would get a logic that is, I believe, not yet cataloged. It would be strange to call it anything other than *dual minimal logic* (cf. Sect. 5.3). A Kripke semantics for dual minimal logic is easily given by dropping the requirement that $\top$ should receive value 1 at every world.

The most significant principle dual minimal logic lacks is the LEM. However, we have seen that the LEM is quite unobjectionable for our purposes. Indeed, to my mind, it is quite a bit less problematic than Explosion is for the verificationistic case.

Instead of the half-cocked notion of $\top$ that leads us to dual minimal logic, we should thus try to secure a $\top$ that will give us full dual intuitionistic logic. I see no conceivable reason why the outcome of that endeavor should be any different from the verificationistic upshot: We will need verifications in the ingredient sense, and even then the required conversions for dual intuitionistic negation will rely on *empty promise* conversions (again, cf. Sect. 5.3).

7.7 Chapter Summary

We looked at dual intuitionistic logic in some of its variations. We were able to model the progress of a purely falsificationistic inquiry using models in which the value 0 instead of the value 1 projected into the future. Consequence was defined as the transmission of unfalsifiability, which here is just what value 1 denotes (but we noted that this will change from the next chapter on, where 1 will mean "verifiable" once again and 0 will still stand for "falsifiable").

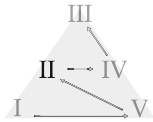
Then, we worked our way through the connectives, with mixed results. While conjunction and disjunction posed no serious problems, the conditional and the negation were disappointments. Dummett's $\rightarrow$, the modus tollens device $\supset_{Tol}$, the intuitionistic conditional $\supset$, and the co-implication $\prec$ all turned out to be unsuitable analyses for indicative conditionals. But such an analysis of items occurring in natural language is surely Dummett's aim, not the study of connectives that only have clear algebraic sense.

Both conditional and negation call for the introduction of verifications into the ingredient sense. We will therefore now leave the bottom floor of the pyramid and move on to the expanded theories. I will first discuss the verificationistic variant, Stage II, and then turn to expanded falsificationism in the following chapter.

[21] Priest (2009), p. 189.

Chapter 8
Stage Two: Expanded Verificationism and the Logic N_3

8.1 Chapter Overview



Up to now, we have looked at pure verificationism and pure falsificationism. However, both pure theories have left us unsatisfied when it came to giving an account of the connectives. Therefore, we are moving up to the second level of the pyramid, in which we will be able to employ both verifications and falsifications in the ingredient sense.

Our first item on the "expanded" level, and the topic of this chapter, is the expanded verificationism of Stage II. To make a correct assertion is once again to say something verifiable. But to know how a statement will behave in complex statements will involve knowing both its verification conditions and its falsification conditions.

As I showed in Chap. 5, we are still following Dummett's footsteps at this stage. However, we found no indication that he thought that this move to an expanded verificationism would have any effect on the logic the constructivist should endorse, i.e., that anything but intuitionistic logic should supply the correct rules of inference. Only when verifications are completely thrown out did he consider changes in the constructive logic, as we have seen in the last chapter.

As there is little in Dummett's texts on how to set up the logic for Stage II, I will simply give the most plausible picture I can come up with. The resulting logic will not be intuitionistic, but rather a form of *Nelson logic*.

I will begin the chapter by going straight into the BHK clauses of this logic. It is actually quite natural to formulate verification and falsification conditions for complex statements, especially as we have the experience of the pure theories behind us and are able to draw on the ideas that worked at those stages. Even some of the ideas that failed us there will now come useful, because now we have the necessary resources in the ingredient sense to make them work.

Then, we will look at the Kripke semantics of the logic. Some thought has to go into the question whether we want to allow for gluts in the semantics, but in the end, I will only allow gaps. That is to say, there will be worlds in which some statements are neither verifiable nor falsifiable, but no worlds in which a statement is both verifiable *and* falsifiable.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_8,

The end of the chapter will bring a bit more discussion of the connectives that we have often found to be in need of comment before: the conditional and the negation.

8.2 BHK Interpretation

Although this chapter is very similar to the last one in structure, I will change the order of things a bit. I will start out by giving the BHK clauses for a Stage II logic, and then, I will present a matching Kripke semantics.

The task of the BHK interpretation is to give both verification and falsification conditions for all of the connectives, so our list of clauses will be twice as long as those we have seen in the lower, pure stages.

Luckily, the task is already almost accomplished, as we have seen all the important ideas in the chapters before. We only need to collect them in one place. Here it goes:

Both conjunction and disjunction are as well behaved as they have always been. To verify a conjunction, verify both conjuncts; for a disjunction, verify at least one of them. As we know from the last chapter, to *falsify* a conjunction, we have to falsify only one of the conjuncts and to falsify a disjunction, we have to falsify both disjuncts.

Also from the last chapter (Sect. 7.6.2), we know what we really want to say about the falsification condition of a conditional: We will need a verification of the antecedent and a falsification of the consequent. To falsify my "If it rains, I'll write my report," you will have to show that it is raining and that I am not writing my report.

The difference to what we found in the last chapter is that now, *we can actually use this definition*, because now we may use both verifications and falsifications in the ingredient sense.

But what of the verification condition for the conditional? Nothing *quite* as perfect as this falsification condition seems on offer, but until now we were doing all right with the intuitionistic understanding: To verify a conditional is to show how any verification of the antecedent can be turned into a verification of the consequent.

Lastly, the negation I want to propose is, as I had mentioned earlier (Sect. 5.3.1), a toggle negation[1] that brings us from verification to falsification and vice versa. To me, this sounds intuitively right, and it exhibits a pleasing symmetry. Moreover, if we have to countenance verifications and falsifications in our semantics, it would seem plausible to suppose that language has a means of going back and forth between the two notions and that this device should be negation.

I will show you the characteristics of the logic that is based on this understanding of the negation and then, at the end of the chapter, come back to the question why we should prefer the toggle negation over the intuitionistic one.

Here, then, are the BHK clauses corresponding to the ideas above:

[1] Symbolized as $-A$.

- c is a verification of $A \wedge B$ iff c is a pair $(c1, c2)$ such that $c1$ is a verification of A and $c2$ is a verification of B
- c is a falsification of $A \wedge B$ iff c is a pair $(i, c1)$ such that $i = 0$ and $c1$ is a falsification of A or $i = 1$ and $c1$ is a falsification of B

- c is a verification of $A \vee B$ iff c is a pair $(i, c1)$ such that $i = 0$ and $c1$ is a verification of A or $i = 1$ and $c1$ is a verification of B
- c is a falsification of $A \vee B$ iff c is a pair $(c1, c2)$ such that $c1$ is a falsification of A and $c2$ is a falsification of B

- c is a verification of $A \supset B$ iff c is a procedure that converts each verification d of A into a verification $c(d)$ of B
- c is a falsification of $A \supset B$ iff c is a pair $(c1, c2)$ such that $c1$ is a verification of A and $c2$ is a falsification of B

- c is a verification of $-A$ iff c is a falsification of A
- c is a falsification of $-A$ iff c is a verification of A.

8.3 Kripke Semantics

Lopez-Escobar and Wansing[2] have argued that the above BHK clauses correspond not to classical logic, but rather to one of the *Nelson logics*, named after D. Nelson (1949).[3] There are two main variants of these logics called N_3 and N_4. N_3 allows for gaps in the semantics, and N_4 features both gaps and gluts. I will introduce Kripke semantics for both of these logics, although eventually I will endorse N_3. I will justify this choice in Sect. 8.4.

Now then: A model for N_3 is once again a structure $[W, \leq, v]$, w being a non-empty set of partially ordered ($\leq$) worlds and v a valuation function from formulas to 1 and 0. Worlds are again intuitively to be understood as stages of investigation, and the accessibility relation marks that one stage is an epistemically possible development from one stage to another.

This time, though, we give both of the values 1 and 0 a substantive reading: 1 stands for "verifiable," 0 for "falsifiable." This is in contrast to the semantics of intuitionistic and dual intuitionistic logic, in which one of the values marked the constructive notion and the other the mere *absence* of that notion.

Moreover, for N_3, we allow v to be a *partial* function, so that statements might not receive either value at a given world. This reflects the fact that at a stage of investigation, a statement might be neither verifiable nor falsifiable. Note that $w \Vdash_0 p$ is not equivalent to $w \nVdash_1 p$ any more and that the same of course goes for $w \Vdash_1 p$ and $w \nVdash_0 p$.

---

[2] Lopez-Escobar (1972); Wansing (1993).

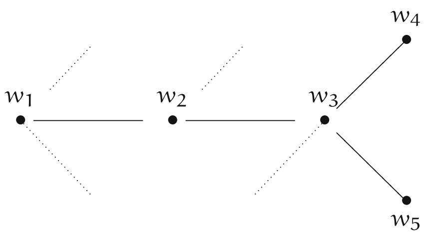
[3] The logic was independently introduced by von Kutschera (1969). For further information on Nelson logic, see Wansing (1993, 1998); Odintsov (2008).

The logic N_4 gives even more options: It allows v to assign 1, 0, neither, or *both* values to a statement at a world. That is, we are not dealing with a valuation function any more, but with a valuation *relation*. This is the only difference between the two logics, and everything else that follows in this section applies to both of them.

We assume that verifications and falsifications are conclusive, and therefore, we will have hereditary constraints for both 1 and 0:

For all p and all worlds w and w', if $w \leq w'$ and $w \Vdash_1 p$, then $w' \Vdash_1 p$, and
for all p and all worlds w and w', if $w \leq w'$ and $w \Vdash_0 p$, then $w' \Vdash_0 p$.

To illustrate, here is the good old example of a Kripke model again:



Here is a valuation that suits the requirements of N_3:

	w_1	w_2	w_3	w_4	w_5
p	-	-	1	1	1
q	-	0	0	0	0
r	1	1	1	1	1
s	-	-	-	1	0

Both 1 and 0 project forward, because both of them record a constructive achievement (verification and falsification, respectively) that is taken to be permanent. There is a third option, here represented by "-": a gap, a mere absence of either verification or falsification. Other than that, all behaves quite as you would expect.

Now, a different valuation will show the peculiarities of N_4: As we said, this logic allows for gluts as well as gaps. The valuation below reflects this by assigning both values, 1 and 0, to some of the statements.

	w_1	w_2	w_3	w_4	w_5
p	-	-	1	1	1
q	-	0	0	0	1,0
r	1	1	1,0	1,0	1,0
s	-	-	-	1	0

Once again, both 1 and 0 project forward, so that a statement that receives both 1 and 0 will never change in status. On the other hand, a statement that is only verified (falsified) may always become both verified *and* falsified later on.

As we are concerned with a species of verificationistic logic, we choose the definition of logical consequence we had given for intuitionistic logic:

$\Gamma \vDash A$ iff in every model and every $w \in w$, if $w \Vdash_1 B$
for every $B \in \Gamma$, then $w \Vdash_1 A$.

8.3.1 The Connectives

We now have to give separate clauses for $\Vdash_1$ and $\Vdash_0$ when defining the connectives. Guided by the above BHK-style clauses, we get:

$w \Vdash_1 A \wedge B$ iff $w \Vdash_1 A$ and $w \Vdash_1 B$
$w \Vdash_0 A \wedge B$ iff $w \Vdash_0 A$ or $w \Vdash_0 B$

$w \Vdash_1 A \vee B$ iff $w \Vdash_1 A$ or $w \Vdash_1 B$
$w \Vdash_0 A \vee B$ iff $w \Vdash_0 A$ and $w \Vdash_0 B$

$w \Vdash_1 -A$ iff $w \Vdash_0 A$
$w \Vdash_0 -A$ iff $w \Vdash_1 A$

$w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$[4]

$w \Vdash_0 A \supset B$ iff $w \Vdash_1 A$ and $w \Vdash_0 B$

---

[4] Some care is needed in transferring the condition from intuitionistic logic to Nelson logic. The clause for intuitionistic logic is sometimes given homophonically as

(1) $w \Vdash_1 A \supset B$ iff for all $x \geq w$, if $x \Vdash_1 A$, then $x \Vdash_1 B$

and sometimes in terms of a disjunction, the way we have done above. Here it is again:

(2) $w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \Vdash_0 A$ or $x \Vdash_1 B$

Given the classical metalanguage and the fact that in intuitionistic logic $x \nVdash_1 A$ is equivalent to $x \Vdash_0 A$, the two clauses are equivalent to

(3) $w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$

In Nelson logic, this equivalence breaks down, because $x \nVdash_1 A$ is *not* equivalent to $x \Vdash_0 A$. The homophonic clause (1) is equivalent only to the last condition, (3), the one I gave above for Nelson logic.

Why not consider (2), though? First, because it simply does not capture the thought that B is verified whenever A is verified, but rather the thought that B is verified whenever A is unfalsified, which does not seem promising as an account of a conditional. Second, the hereditary conditions ensure that this is actually equivalent to

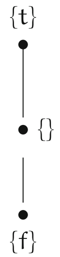
(4) $w \Vdash_1 A \supset B$ iff $w \Vdash_0 A$ or $w \Vdash_1 B$

Choosing this clause then makes the only intentional notion in the definitions superfluous, and all our worlds and heredities come to nought: As will become apparent presently, we would just have defined K3 or FDE with a material conditional in a complicated way.

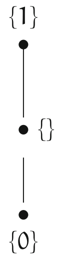
Conjunction, disjunction, and negation look pleasantly unspectacular. Negation is, unlike intuitionistic negation, an extensional connective that only concerns itself with the world at hand.

Indeed, note that the only clause that makes reference to other worlds at all is the positive clause for the conditional. The first-degree (conditional-less) fragments of N3 and N4 are actually equivalent to two systems we have met before: K3 and FDE. In the Dunn style semantics we met on page 4.3.2, the connection is quite clear:

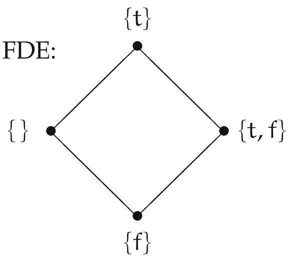
K3:



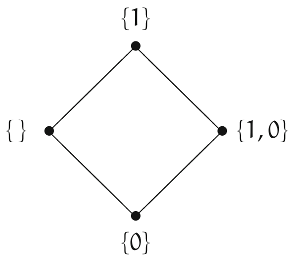
N3:



FDE:



N4:



1 obviously corresponds to t and 0 to f ; the only real difference lies in the interpretation of these values. Whereas before, we heard talk of "told-truth values," we now want to read these values as "verifiable" and "falsifiable."

Given all this, it is not surprising that one thing that clearly marks a difference between N3 and N4 is that N4 is a paraconsistent logic. N3, on the other hand, will never give us a situation in which both $w \Vdash_1 A$ and $w \Vdash_1 -A$ hold, and thus, the consequence relation becomes explosive.

I will talk more of characteristic features of Nelson logic, and in particular about the conditional, below. But first, it is time to make a choice between N3 and N4 for our project.

8.4 Do We Want Gluts?

So, should we want to use models of the N3 variety, in which no glutty valuations are allowed, or move to the more liberal N4 models, in which gluts are available?

The answer to this question hinges on whether we want to allow for statements that are both verifiable and falsifiable.

We have an uncharacteristically clear notion of what Dummett's own answer to this question is: He explicitly rejects gluts, as we have seen in the quote on p. 91. Nothing that is verified can ever be falsified, nothing that is falsified can be verified. That is, to reconstruct his vision of a Stage II theory, we should certainly stay away from glutty models.

There are other voices around, though. To see their point, we have to come back to the question we have bracketed up to now: Are we really well advised to model our semantics on "conclusive" verifications and falsifications?

The case for gluts is based on the fact that, in the empirical realm, we must always be prepared to find our best corroborated hypotheses to fail, and to being forced to accept what we had formerly thought was ruled out by the evidence. Verifications and falsifications are based on evidence, and evidence is usually taken to be defeasible.

This line of argument is pursued by Cogburn (2004). His considerations are directly aimed at Dummett's project. He argues that verificationists of Dummett's ilk should embrace not only gluts in the semantics, but also an outright dialetheism.

The central claim of the paper can be summarized thus: Once we move out of the pristine realm of mathematics and step into the rough and dirty empirical world, there is no such thing as absolutely certain and conclusive proof any more. The most we may hope for is very good but still defeasible evidence. But then, it is hard to deny that there might be situations where we are confronted with very good evidence for a statement A and equally good evidence for its negation $\neg A$.

Then, if verifiability is spelled out in terms of very good evidence, we must face the possibility of verified contradictory statements. If, furthermore, negation is the kind of verification–falsification switch that we have taken it to be in this chapter,[5] we end up with a glutty semantics. And, lastly (and optionally), if we want to make the final leap from verification to truth, we end up with genuine dialetheism.

Now, Cogburn is well aware of the fact that there is a rather obvious response to his line of argument: If the evidence we are dealing with here is supposed to be defeasible, why not say that, if we have good evidence for both a statement A and its negation $\neg A$ (or, equivalently for our purposes, verifying and falsifying evidence for the same statement A), these bodies of evidence cancel each other out, so that neither A nor $\neg A$ is verified?[6]

There are several answers Cogburn gives to this, not all of them bear directly on our topic. For our purposes, the most interesting of these answers is the last one:

> As a final note, the claim that a warrant for P automatically undermines warrants for $\neg P$ involves an ugly equivocation between not having enough evidence for a claim and having too much evidence for the claim and its negation. For the dialetheist, this is obviously not

[5] Although Cogburn does not point to any specific logical systems that might be candidate companions to his view, I think that N_4 would be well up to the job.

[6] This is the question that Cogburn addresses. Maybe he is going a bit too easy on himself here, for the question could be put more polemically as: "If all evidence is supposed to be defeasible, what *else* could defeat good evidence for A, if not good evidence for $\neg A$?".

so. When there is not enough evidence, we can either optimistically wait and hope, or we can decide the claim is neither true nor false. When there is too much evidence both ways, the claim is both true and false.

Indeed, I believe that he is right: The distinction that is in danger of being smeared over is one that we should like to keep making. It is the distinction between a just-so story and a carefully researched report that acknowledges contrary evidence.

However, I also believe that the intuition that contrary evidence can disqualify a verification is worth holding on to. In principle, the Nelson models give us enough fine structure to accommodate both of these demands.

The proposal would be as follows: A verification is not only a body of very good evidence, but rather good evidence *plus the absence of good counterevidence*.

At this point, you should be reminded of the *exactly true logic* ETL that I introduced in Sect. 4.4. The logic that is beckoning, call it $ETL_{\supset}$, is related to ETL in the same way as N_4 is to FDE. The difference is effected by an obvious modification of the consequence relation of N_4:

Consequence $ETL_{\supset}$:

$\Gamma \vDash A$ iff in every model and every $w \in w$, if $w \Vdash_1 B$ and $w \nVdash_0 B$ for every $B \in \Gamma$, then $w \Vdash_1 A$ and $w \nVdash_0 A$.

As one would expect, the rather strange syntactic features of ETL are not alleviated by adding a conditional to it. I will not delve into the resulting system, for I believe that this proposal, while conceptually an improvement on N_4 for this application, is actually not going far enough.

Consider the case of a statement for which we have evidence that would be good enough to verify it, were it not for the additional good evidence we have for its negation. In the semantics for $ETL_{\supset}$, such a statement could never become assertible, due to the heredity constraint for the two semantic values. However, it seems that the idea of defeasible evidence entails not only that a statement taken to be verified might become unverified due to counterevidence, but also that this counterevidence might become so overwhelmingly strong that the original evidence becomes negligible and the negation of the statement will be considered verified.

If *this* kind of dynamic is to be modeled, then we need to turn our attention to constructive non-monotonic systems. Incidentally, I believe this is one of the most interesting ways in which the ideas in this book could be further developed, but the topic is too large for me to attempt even a start of an investigation. For the rest of these pages, I will keep assuming the conclusiveness of verifications and falsifications, and if this assumption turns out to be unacceptable, then at least I think some groundwork has been laid on which new non-monotonic systems can be based.[7]

---

[7] One finds a non-monotonic system based on Nelson models on p. 144 of Wansing (1998). Some further, very promising ideas in this direction are in Jaspars (1994, 1995). Yet another suggestion, made by Johannes Marti in discussion, is to simply do away with the heredity constraints altogether. While this would indeed ensure a great deal of revisability (maybe even too much), unnegated conditionals would still stay true throughout. It does not seem that this is a promising route to take.

8.5 Features of the Logic N_3

If we assume that verifications and falsifications are conclusive, then I can see no reason to hold on to N_4, so I will concentrate on N_3 and its characteristic features from now on.

I will contrast these features with intuitionistic logic, the logic that Dummett seemed to think apt for a Stage II theory. As has been hinted at before, all intuitionistic inferences can be mimicked in N_3, because intuitionistic negation is actually definable in the following way:

$$\sim A =_{\text{def}} A \supset -A$$

With the aid of the Kripke clauses, it is easy to check that the formula $A \supset -A$ will be verified iff A is never verified in a subsequent world. It is also clear that $\models -A \supset\sim A$, but not $\models\sim A \supset -A$. For this reason, the Nelson negation is also known as *strong negation*.[8]

With all that in mind, here now is a list of similarities and differences we find with regard to intuitionistic logic:

	IL	N_3
$--A \models A$	n	Y
$A \models --A$	Y	Y
$\models A \vee -A$	n	n
$-(A \vee -A) \models B$	Y	Y
$\models -(A \wedge -A)$	Y	n
$(A \wedge -A) \models B$	Y	Y
$-(A \wedge B) \models (-A \vee -B)$	n	Y
$A, A \supset B \models B$	Y	Y
$-B, A \supset B \models -A$	Y	n
$A \supset B \models -B \supset -A$	Y	n

(Footnote 7 continued)
Note that alongside a non-monotonic system for verifications and falsifications, we could keep using N_4 to keep track of information that supports a statement, with the understanding that such information need not amount to a verification. However, even a piece of information that will be defeated as evidence is a piece of information that we have and that we keep. Therefore, the irreversibility of "support for A and support against A" that results from the heredity constraint in N_4 would make perfectly good sense here. For an interpretation of N_4 that lays more emphasis on such an informational interpretation, see Ref. Wansing (1993).

[8] At least this is the genealogy Vakarelov (2006) offers for the name. A reasonable choice of terminology on my part might have been to call every negation I call "toggle negation," i.e., every negation that switches between verification and falsification, "strong negation." I have chosen not to because (a) the definability of intuitionistic negation as well as the validity of $\models -A \supset\sim A$ depends on the conditional as well as the negation, and toggle negation might appear in a system which is missing these features, (b) because there are other uses of *strong negation* in other parts of the logical literature, and (c) because the name "toggle negation" seems to me more suggestive of the essential role it plays in a verification–falsification semantics.

We find, once again, that the toggle definition of negation gives us *both* double negation laws. LEM fails, as in intuitionistic logic: Not every statement is either verifiable or falsifiable. However, $-(A \vee -A)$ can never be verified, so it implies any arbitrary statement, again just as it does in intuitionistic logic.

But in Nelson logic, as is clear from the validity of Double Negation Elimination, the Law of Excluded Third, $--(A \vee -A)$, must fail to be valid as well. This marks a difference to intuitionistic logic, as we had seen in Chap. 3.

Semantically, these validities seem to have the correlations that Dummett predicted: Bivalence clearly fails, but also *tertium non datur* must go: We are unashamedly endorsing gaps in the semantics of N_3.

Even absolutely undecidable statements, statements which we know will never be decided, are a possibility in Nelson logic. The formula $(A \supset -A) \wedge (-A \supset A)$[9] is satisfiable, but only in a model in which A fails to receive a value at all worlds. This allows the constructivist employing N_3 to say something that the intuitionist could not, for to know that A would never be verified is for the intuitionist already enough to prove $\sim A$.

The failure of $\vDash -(A \wedge -A)$ corresponds to the failure of LEM, and it is instructive to see how close the connection is. $-(A \wedge -A)$ will be true if there is a falsification of $(A \wedge -A)$. Such a falsification, under the present proposal, will consist in either a falsification of A or a falsification of $-A$. It would be just as preposterous to claim to be able to be able to supply either a falsification for every statement or its negation as it was preposterous to claim to always be able to verify one of them. Therefore, both $\vDash (A \vee -A)$ and $\vDash -(A \wedge -A)$ will have to go.

We saw in Sect. 3.3 that there was a more general meta-logical property of intuitionistic logic connected to the failure of LEM, the *disjunctive property*: A disjunction is a theorem of a theory closed under intuitionistic logic iff at least one of the disjuncts is a theorem as well.

N_3 has this property as well, but in addition, it has the constructible falsity property [10]: A negated conjunction, $-(A \wedge B)$, is a theorem of a theory closed under N_3 iff either $-A$ or $-B$ is a theorem.

But of course, all this does not mean that $(A \wedge -A)$ can be *verified* in a semantics without gluts. Indeed, it obviously cannot be verified, and as a consequence, Explosion, $(A \wedge -A) \vDash B$, is valid. As I said above, this marks a principal difference between N_3 and N_4.

$-(A \wedge B) \vDash (-A \vee -B)$, along with all of the other de Morgan's laws, is valid. Spelling out the verification conditions of antecedent and consequent makes it quite clear that both are verifiable in the same circumstances: when at least one of A or B is falsifiable. There is nothing non-constructive inherent in this inference.

So far, none of the features of N_3 seem objectionable. The same goes for the next line in the list above: Modus ponens holds in N_3, as we surely should have hoped.

On the other hand, the next two items seem less appealing. Both contraposition $(A \supset B \vDash -B \supset -A)$ and modus tollens $(-B, A \supset B \models -A)$ fail in N_3. The

[9] Or, with the definition of intuitionistic negation above in place, $\sim A \wedge \sim -A$.

[10] Cf. Wansing (2008), p. 342.

counter model to both of these inferences is extremely simple: A model with only one world at which B is falsifiable, but A is neither verifiable nor falsifiable.

Informally, the reason why the two principles fail is that not much is assumed about the relation of verifications and falsifications (other than that no statement can be both verified and falsified). If you know how to turn a verification for A into another verification of B, then the logic does not want to make a commitment to your being able to turn a falsification of B into a falsification of A.

Interestingly, seeing how very distinctive a feature this failure is, I have seen little by way of examples that show that verifications and falsifications are indeed so unrelated. A simple-minded example would be this:

If you can verify that the test slip turns blue after you put it into the test-tube, you can thereby verify the presence of substance XYZ in the tube. This does, according to Nelson logic, not imply that if you have some unrelated means of falsifying the presence of XYZ in the tube, you can from that construct a method to falsify the claim that the slip would turn blue if you put it into the tube. But, at least in this example, it seems clear that you can come up with a conclusive method of falsifying the claim: Just put the test slip into the tube!

Not that there might not be more sophisticated counter examples, but intuitively, one might well expect contraposition and modus tollens to hold for a conditional that is explained in terms of verifications and falsifications.

8.5.1 Some Attempts to Get Contraposition Back

We can of course ask whether we can find a different positive clause for the conditional, one that gives us contraposition and modus tollens back. In Sect. 7.6.2, we had met $\supset_{\text{TOL}}$, the conditional that was based on the following BHK clause:

- c is a verification of $A \supset B$ iff c is a procedure that converts each falsification of B into a falsification of A

Back then, we had little use for the idea, because we were not allowing ourselves talk of verifications. Here, of course, we feel no such inhibition, and we proceed to consider the Kripke clause:

$$w \Vdash_1 A \supset B \text{ iff for all } x \geq w, x \nVdash_0 B \text{ or } x \Vdash_0 A$$

I will adopt the name $N3_{TOL}$ for the logic that we get by using this clause instead of the standard one. It would be rather surprising if this clause did not bring us modus tollens back, and indeed, it does. However, contraposition still fails, and in exchange for modus tollens, we have to give up modus ponens.[11] Not a great improvement, then.

[11] Counter model for both $A \supset B \vDash -B \supset -A$ and $A, A \supset B \models B$: A model with one world at which A is verified and B neither verified nor falsified.

Seeing that neither the intuitionistic idea of a transformation of verifications nor the alternative idea of transforming falsifications works too well, we might also consider using *both* of these conditions simultaneously.

Two alternatives present themselves here: First, we might say that a conditional is verified if one or the other of the conditions hold:

$w \Vdash_1 A \supset_{OR} B$ iff for all $x \geq w$, $x \nVdash_0 B$ or $x \Vdash_0 A$ OR for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$

Call the logic that results from adopting this condition N_{OR}.

Second, we can be more demanding and ask that *both* conditions should have to hold:

$w \Vdash_1 A \supset_{AND} B$ iff for all $x \geq w$, $x \nVdash_0 B$ or $x \Vdash_0 A$ AND for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$

I will call the Nelson logic that features this conditional N_{AND}.[12]

Here is a list of the salient differences between these options:

	N_3	N_{3TOL}	N_{OR}	N_{AND}
$A \supset B \models -B \supset -A$	n	n	Y	Y
$-A \models A \supset B$	Y	Y	Y	n
$B \models A \supset B$	Y	Y	Y	n
$A, A \supset B \models B$	Y	n	n	Y
$-B, A \supset B \models -A$	n	Y	n	Y

Of the four alternatives, we seem to get by far the most pleasing results from N_{AND}. Not only contraposition, modus ponens, and modus tollens all hold up, we are getting rid of $-A \models A \supset_{AND} B$ and $B \models A \supset_{AND} B$, two inferences that were never too much in favor anyway.

However, there is a price to pay: The deduction theorem will fail us. There will be cases in which we have $A \models B$, but not $\models A \supset_{AND} B$. The reason for this is that the failure of contraposition is built into the turnstile: We may have $A \models B$, but not $-B \models -A$.[13] A logic in which the conditional contraposes, but the deduction theorem fails and contraposition around the turnstile fails seems a lackluster affair.

Now, we could try to make everything right by tweaking the definition of logical consequence we gave. Instead of only verifiability preservation left to right, we might additionally require falsifiability transmission right to left:

Contraposable Consequence:

$\Gamma \models_{Contrap} A$ iff in every model and every $w \in w$, if $w \Vdash_1 B$ for every $B \in \Gamma$, then $w \Vdash_1 A$, and if $w \Vdash_0 A$, then $w \Vdash_0 B$ for some $B \in \Gamma$

[12] $\supset_{AND}$ is mentioned in Chap. 12 of Rasiowa (1974), albeit in a different framework, where it is used alongside the original Nelson conditional.

[13] A related phenomenon is that logical equivalents may not be substituted freely in formulas; only when also their negations are logically equivalent this will be possible. For example, double negations may be introduced and eliminated inside formulas, as both A and $--A$ on the one hand, and on the other hand $-A$ and $---A$ are logically equivalent.

This consequence, together with $\supset_{AND}$, indeed gives us all the contraposition we could hope for. But the loss is, once again, grave: Neither modus tollens nor modus ponens hold any more.[14]

It is from the frying pan into the fire, and then from there into something else that is at least equally unpleasant.[15]

8.5.2 *Embracing Contraposition Failure*

As none of these options seem *much* better than the standard clause for the Nelson conditional, maybe we should simply make our peace with the idea that we have to do without contraposition and modus tollens. And maybe we should look more carefully for counterexamples.

In ending up with a logic in which contraposition fails, we are actually in good company. Both R. Stalnaker and D. Lewis developed so-called *conditional logics.* In these logics, contraposition fails, but modus tollens (unlike in N_3) holds. Both authors bite the bullet and argue that contraposition is invalid for natural language conditionals.

Here is the counterexample Stalnaker gives:

> For an example in support of this conclusion [i.e., that contraposition is invalid], we take another item from the political opinion survey: "If the U.S. halts the bombing, then North Vietnam will not agree to negotiate." A person would believe that this statement is true if he thought that the North Vietnamese were determine to press for a complete withdrawal of U.S. troops. But he would surely deny the contrapositive. "If North Vietnam agrees to negotiate, then the U.S. will not have halted the bombing." He would believe that a halt in the bombing, and much more, is required to bring the North Vietnamese to the negotiating table (Stalnaker 1968, p. 107).

If you are like me, then you will have to read this twice before you fully understand what is supposed to be going on here. It is much easier to understand if you add a small word at the beginning of the original conditional:

> *Even* if the U.S. halts the bombing, then North Vietnam will not agree to negotiate.

Lewis gives a different example that is likewise helped along by a small word, in this case "still." He asks us to consider the inference:

> If Boris had gone to the party, Olga would still have gone.
>
> ∴ If Olga had not gone, Boris would still not have gone (Lewis 1973, p. 34).

He goes on to argue that this is actually an invalid inference: "Suppose that Boris wanted to go, but stayed away solely in order to avoid Olga, so the conclusion is false; but Olga would have gone all the more willingly if Boris had been there, so the premise is true."

---

[14] The counter model for modus ponens is again the one-world model in which B is falsified and A a gap, the counter model for modus tollens the model in which A is verified and B a gap.

[15] Appendix 1 collects some more information on N_{TOL}, N_{AND} and N_{OR}.

This is easier to follow than Stalnaker's original example, but cross out the "still," and it becomes much harder to deny that the inference seems quite compelling. I have to say that I have some doubts that the conditionals in these examples really are normal conditionals. It seems that an "even if... then..." conditional calls for a different analysis than a simple "if... then..." conditional.[16]

In a later work, Stalnaker addresses this worry and suggests that the "even" has a purely pragmatic effect on conditionals.[17] It might be pragmatically unacceptable for someone with the political view above to *say* "If the U.S. halts the bombing, then North Vietnam will not agree to negotiate." However, if asked to judge the truth of it, she will agree to it.

If these conditionals really *are* ordinary conditionals that show the invalidity of contraposition in natural language, then the friend of N_3 has reason to smirk. For these, examples motivate her logic better than the logics of Lewis and Stalnaker, because they *also* invalidate modus tollens.

Assume that "If the U.S. halts the bombing, then North Vietnam will not agree to negotiate." is true for the reason above and that North Vietnam in fact agrees to negotiate. You should better not be forced to draw the inference that the U.S. did not halt the bombing from this.

But this inference should go through in Stalnaker's and Lewis's systems.[18] N_3, on the other hand, has an easy time telling you why the inference should not be drawn: It is invalid.

So far, so good. However, the reason why contrapostions fail in the Lewis-Stalnaker examples on the one hand and in N_3 as interpreted by the BHK clauses on the other hand do not seem to match up too well. The conditional "If the U.S. halts the bombing, then North Vietnam will not agree to negotiate." is, if it is true, not true because *every* verification of the antecedent can be transformed into a verification of the consequent. To verify that the USA halts the bombing and withdraws all its troops is, in part, to verify the antecedent, but in this case, the consequent will not be verifiable.

It would seem that to do full justice to these examples, the Nelson conditional would have to be supplied with some mechanism to capture *ceteris paribus* (or "all other things being equal") clauses. This is yet another interesting task that I will not attempt to go into here.

Here is an example that is closer to home. One of the consequences of our setup is that

If A, then it is verifiable that A.

is always verifiable. For, if A is verifiable, then surely it is verifiable that A is verifiable. There is just no good reason to deny this. However, the contrapositive

If it is not verifiable that A, then not A.

---

16 This view of the matter is defended in McLaughlin (1990), though the analysis of the "even if" conditionals is not pursued very far.

17 Stalnaker (1984), p. 124.

18 It is possible to modify these systems to get rid of modus tollens, though.

need not be verifiable. The antecedent must be strengthened to “It is falsifiable that A” for this conditional to go through. I trust this assessment of the second conditional is relatively uncontroversial. What might be less uncontroversial is that we should want to endorse each and every instance of the original conditional. In particular, if verifiability is identified with truth, some constructivist might feel a bit uncomfortable with it. Is what is true at this moment really only that which is verifiable at this moment?

For my part, I think they should not feel embarrassed by this consequence. It might violate some intuitions, but this is better than to patch up the view in some ad hoc way to cater to those intuitions.

If they feel this consequence is really unacceptable, then maybe they should consider an eclectic or pluralistic view on logic, such as sketched earlier on in Sect. 2.11. I will come back to this topic in the last chapter.

However, a different way to avoid commitment to this conditional is to interpret it as $\supset_{\mathrm{AND}}$. In that case, it will fail to hold precisely because its contrapositive holds. But then, one will be denying that every instance of “If A, then it is verifiable that A.” is verifiable, but conceding that “It is verifiable that A.” follows logically from “A”.[19] Again, I would feel somewhat uncomfortable with this result, but maybe it can be made plausible after all.

To sum up these considerations about the conditional, I will (somewhat tentatively) keep endorsing the original account we find in N_3, viz., the positive clause of the intuitionistic conditional, and with it the failure of contraposition and modus tollens. Of the alternatives I considered, $\mathrm{N}_{\mathrm{AND}}$ seems by far the most tempting logic, and I will be keeping an eye on it in what follows. To endorse $\mathrm{N}_{\mathrm{AND}}$ would suggest to argue (1) that the deduction theorem is not as essential as it is usually taken to be and (2) that “even if” conditionals are not really normal conditionals with some pragmatic swirls added, but rather a whole different type of conditional.

8.6 Toggle Negation Versus Intuitionistic Negation

Let me end this chapter by picking up the question that was raised back in Sect. 5.3.1. Given the choice between intuitionistic negation and toggle negation, which should we go for? As I said repeatedly, I find the toggle negation in Nelson logic much more natural. Of course, an intuitionist may just shrug his shoulders at this and claim that his intuitions point elsewhere. Let me then try to give some more substantial argument for my preference.

One tempting line of argument would be this: As we had seen, the intuitionistic negation depends on our acceptance of *empty promise* conversions (cf. p. 39). The toggle negation, on the other hand, is not dependent on this idea, so in order to attack intuitionistic logic, one might argue that such conversions are inadmissible.

[19] Unless one also goes on to embrace the modification of logical consequence that is $\vDash_{\mathrm{Contrap}}$.

However, we find similar empty promise conversions in N_3 as well. Just consider the valid inference

$$-A \models A \supset B$$

To make a strong argument against empty promise conversions while accepting such an inference seems to make for an unstable position.

Of course, if there in fact *is* a problem with empty promise conversions and $-A \models A \supset B$, then it should be pointed out that it is not really a problem with the toggle negation in Nelson logic. The problem lies in the positive clause for the conditional (which of course is the same as in intuitionistic logic), just as the problem with intuitionistic negation can be seen as a problem of the conditional if we think of $\sim A$ as defined as $A \supset \bot$.

Still, if the contest is not between toggle negation and intuitionistic negation in isolation, but between N_3 and intuitionistic logic at large, then both seem committed to empty promise transformations. If these are unacceptable from the constructive point of view, then the Nelson logician should look for a new verification clause for his conditional; again, the conditional $\supset_{AND}$ seems like an interesting alternative here.

If we stick to N_3, however, then we can find nothing to object to in intuitionistic negation. Indeed, as we have seen, it would be strange if we could, because intuitionistic negation, $\sim$A, is definable in N_3[20] as $A \supset -A$!

And this is exactly, I believe, the strong point of Nelson logic. It has no reason to deny intuitionistic negation's legitimacy as a constructive notion, nor does it have to show that it is unnecessary.

On the other hand, the intuitionist who rejects toggle negation has to show either of two things:

Claim 1: Toggle negation is objectionable from the constructive viewpoint. Its definition is somehow flawed.

Claim 2: We don't need toggle negation, and it should fall victim to Occam's Razor.

To seriously make Claim 1 seems absurd. If we do allow verifications and falsifications in the ingredient sense, then the definition of toggle negation seems quite as innocent as the definitions of conjunction and disjunction. As we have just seen, one can envisage doubts about the legitimacy of the conditional, but surely not about the simple verification/falsification switch that is $-$. The only position the intuitionist could take that would support some resistance would be to claim that all ways to show that a verification of A is impossible must suffice to falsify A. In this case, 'It is falsifiable that A is falsifiable' does not imply 'A is verifiable', but this is an implication that has to be accepted along with the toggle negation account. However, I have rejected "A is impossible to verify" as an explication of "A is falsifiable" back in Sect. 6.4.

It would have to be Claim 2, then: We do not need toggle negation, so we should not have it in our logic.

---

[20] And also in N_{AND}.

Now, it seems that the plausibility of this claim depends on the exact project we are engaged in. Are we, for example, trying to give a constructive interpretation of mathematics only? Then, maybe, intuitionistic negation might suffice to express all that we need. If we are, on the other hand, going for a full-scale verificationistic revision of the meaning of negation in all areas of discourse, then the availability of toggle negation is clearly an asset.

If we revise logic, we are in effect claiming that normal people have a somewhat incorrect grasp of the meaning of the logical constants they use. One would suppose that we should be as charitable as possible here. In other words, we should employ a principle of minimum mutilation: The less revision our interpretation makes necessary, the better.

It seems clear that speakers often see no problem in canceling a double negation. But to argue that they always do so would involve some very detailed investigation into negation items in natural language. Is, say,

> She is not unhappy, therefore she is happy

a case of Double Negation Elimination? If so, the intuitionist may have a point in rejecting DNE. On the other hand, the negation items here seem quite different in kind, which might incline us to look for a logical system with more than one negation. The task of finding the right representation of natural language negations is formidable, fascinating, and beyond the scope of this book.[21]

However, there is at least one item in English that surely marks negation: “Not.” And there are cases in which an intuitionistic negation would be verifiable, but a natural language statement featuring “not” would not be judged verifiable. Moreover, toggle negation agrees with natural language intuitions in these cases.

An example is this statement:

> The largest Brachiosaurus alive on September 1st 154888328 B.C. was female.

Suppose that we know that the only way to verify this claim is to make a time travel and that the idea of time travel is absurd. Then, any verification of this statement would lead to the absurd conclusion that time travel is possible. Thus, if the “not” in the following statement is to be analyzed as intuitionistic negation, then this statement is:

> The largest Brachiosaurus alive on September 1st 154888328 B.C. was *not* female.

But the intuitions of most speakers are surely that neither of these examples is verified, and toggle negation explains why in a most straightforward manner: We can neither verify nor falsify these claims.

If there are examples in which a constructive account of the language use of ordinary speakers is only possible if we employ toggle negation as a tool of analysis, then I think we should do so. The alternative would be to argue that these speakers are at fault. But, if I am right about the constructive admissibility of toggle negation, it is quite unclear on what this criticism of actual practice should be based.

[21] The best place to start an exploration of such matters is still, I think, the classic (Horn 1989).

8.7 Chapter Summary

We have met the constructive Nelson logic N_3, a logic that is able to deal with verifications and falsifications side by side. In the semantics, we allow for gaps, but not for gluts. The logical consequence relation transmits, just as in intuitionistic logic, the property of verifiability.

Concerning the connectives, we made the following choices: The device that takes us from verification to falsification and back again is negation, a simple toggle device that needs no information from other worlds in the Kripke semantics. A direct consequence of our adopting this kind of negation was that all the double negation laws are valid. Together with the account of conjunction and disjunction, we came to see that the de Morgan laws are likewise valid, while both LEM and LET failed to be valid.

A peculiarity of the account is that the conditional does not support contraposition, nor is modus tollens a valid rule. However, I gave some arguments for accepting these failures. Admittedly, the matter was not resolved beyond doubt, and a new conditional, $\supset_{AND}$, appeared to be a close contender for the job.

Lastly, I compared N_3's toggle negation and the intuitionistic negation. The upshot of that discussion was this: Even if there might be cases in which we should want to model natural language negations by intuitionistic negations, we have that resource available in N_3, as intuitionistic negation is definable.

Therefore, a supporter of N_3 can't and does not have to complain too much about intuitionistic negation. The only thing he has to show is that his negation can play a useful role in analyzing the meaning of logical constants.

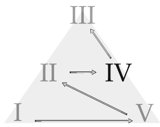
On the other hand, I argued that the concept that toggle negation captures is clearly one of negation, clearly constructive, and clearly useful, and that we should accordingly adopt it.

This summarizes what I have to say about the expanded verificationism of Stage II. In the next chapter, which is on the expanded falsificationism of Stage IV, we will once again ban verifications from the assertoric sense and adopt the falsificationistic stance: As at Stage V, a statement will be correctly assertible iff it is not falsifiable.

However, we will keep employing verifications and falsifications alongside each other in the ingredient sense. The nice consequence of this is that we can just employ the definitions of the connectives that we have worked out in this chapter. The only task is to tweak the definition of logical consequence and to see where that leads us. Unfortunately, some problems lie ahead as well.

Chapter 9
Stage Four: Expanded Falsificationism and the Logic N_{3f}

9.1 Chapter Overview



We are back on the right side of the pyramid. That is to say, we are again interpreting assertibility as unfalsifiability, as we had done back Stage V. The difference is that this time, we can give a better account of the meaning of the complex statements, because we can now make use of verifications in the ingredient sense.
In fact, we can simply reuse the semantic clauses we specified in the last chapter. Characterizing the logic will simply be a matter of adjusting the notion of logical consequence in order for it to transmit unfalsifiability. I will do that and then proceed to inspect the characteristic features of the logic, which I call N_{3f}.

Among these salient features is the fact that N_{3f} is paraconsistent. I argued in Sect. 7.5.1 that if non-falsifiability is what is to be carried from premises to conclusions, we should indeed hope for a paraconsistent logic. The intuitionists suppose that it is sheer vanity to suppose that we could prove every mathematical proposition or its negation and therefore reject the Law of Excluded Middle. The same kind of vanity would be required to make the claim that we could falsify every proposition or its negation. Therefore, we might end up with contradictory premises that are nonetheless not falsifiable. If the logic were not paraconsistent, it would allow the inference from such premises to any proposition whatsoever, even those that are clearly falsifiable, contradicting the assumption that non-falsifiability was a property that was preserved from premises to conclusion.

However, there is a worry here I failed to mention in Chap. 7: We are rightly making space for the possibility of faultless disagreement here, by allowing situations in which A and $\neg A$[1] are both assertible. But those situations will also make $A \wedge \neg A$ unfalsifiable and therefore assertible. Do we really want to allow people to utter contradictions like that and exempt them from criticism, provided that A is neither verifiable nor falsifiable?

[1] Reminder: I'm using $\neg$ as generic negation, in this case to stand in for both $\sim$ and $-$.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_9,

A prima facie very different problem is that modus ponens fails in $N3_f$. This is a quite worrying feature, as this inference is usually assumed to be the very essence of a good conditional.

In the end, these two problems will find a uniform solution. I will introduce the concept of *incoherence*, diagnose both the dubious assertibility of $A \wedge \neg A$ and the failure of modus ponens as grounded in incoherence, and describe a strategy how to deal with these and similar cases.

9.2 Falsificationistic Nelson Logics

Let us transform the logic we found suitable for a Stage II theory into one that satisfies the demands of a Stage IV theory. We will get right to the Kripke semantics. Up to the definition of logical consequence, everything just stays as it was in the last chapter. Incurring obvious redundancy for the sake of easy reference, here is the setup again:

A model is again a structure $[W, \leq, v]$, W a non-empty set of partially ordered ($\leq$) worlds or information states and v a partial valuation function from formulas to 1 (verifiable) and 0 (falsifiable). We also have the hereditary constraints again:

For all p, and all worlds w and w', if $w \leq w'$ and $w \Vdash_1 p$, then $w' \Vdash_1 p$, and
For all p, and all worlds w and w', if $w \leq w'$ and $w \Vdash_0 p$, then $w' \Vdash_0 p$.

We defined the connectives as follows:

$w \Vdash_1 A \wedge B$ iff $w \Vdash_1 A$ and $w \Vdash_1 B$
$w \Vdash_0 A \wedge B$ iff $w \Vdash_0 A$ or $w \Vdash_0 B$
$w \Vdash_1 A \vee B$ iff $w \Vdash_1 A$ or $w \Vdash_1 B$
$w \Vdash_0 A \vee B$ iff $w \Vdash_0 A$ and $w \Vdash_0 B$
$w \Vdash_1 -A$ iff $w \Vdash_0 A$
$w \Vdash_0 -A$ iff $w \Vdash_1 A$
$w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$
$w \Vdash_0 A \supset B$ iff $w \Vdash_1 A$ and $w \Vdash_0 B$

There was a second positive clause for the conditional that we found to be an interesting alternative:

$w \Vdash_1 A \supset_{AND} B$ iff for all $x \geq w$, $x \nVdash_0 B$ or $x \Vdash_0 A$ AND for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$

In N_3, consequence is defined as follows:

$\Gamma \vDash A$ iff in every model and every $w \in W$, if $w \Vdash_1 B$ for every $B \in \Gamma$, then $w \Vdash_1 A$.

This of course means that in N_3, the property that is transmitted from premises to conclusions is verifiability. Because we want to transmit *non-falsifiability,* we define consequence slightly differently:

$\Gamma \vDash A$ iff in every model and every $w \in W$, if $w \nVdash_0 B$ for every $B \in \Gamma$, then $w \nVdash_0 A$.

I call the resulting logic $N3_f$, and in general, I will write X_f for the falsificationistic logic that one gets from a verificationistic logic X by redefining the consequence relation in the above way.

The following table gives an overview over those differences and similarities between $N3$ and $N3_f$ I want to focus on. I'll also give the corresponding differences between N_{AND} and $N_{\text{AND}f}$, although in this chapter nothing much will hang on which positive clause for the conditional we choose, and I will mainly be talking about $N3_f$.

It is clear at a glance that exchanging the property transmitted in logical consequence has yet again a dual effect on the inferences. This time, the easiest way of getting from a valid inference in $N3$ to a valid inference in $N3_f$ is to exchange premise and consequence of a single premise inference and to negate both of them. The same goes for invalid inferences. I'll make some straightforward simplifications, though, that give pithier inferences. For example, to get from $B \vDash -(A \wedge -A)$ to $(A \wedge -A) \vDash B$, I choose $-B$ as an instance of B in the first inference, dualize, and then eliminate double negations on both sides.[2]

	$N3$	$N3_f$	N_{AND}	$N_{\text{AND}f}$
$B \vDash A \vee -A$	n	Y	n	Y
$-(A \vee -A) \vDash B$	Y	n	Y	n
$B \vDash -(A \wedge -A)$	n	Y	n	Y
$(A \wedge -A) \vDash B$	Y	n	Y	n
$A \supset B \vDash -B \supset -A$	n	Y	Y	Y
$-(-B \supset -A) \vDash -(A \supset B)$	Y	n	Y	Y
$-A \vDash A \supset B$	Y	Y	n	Y
$B \vDash A \supset B$	Y	Y	n	Y
$\neg A, A \vee B \vDash B$	Y	n	Y	n
$A, A \supset B \vDash B$	Y	n	Y	n
$-B, A \supset B \vDash -A$	n	n	Y	n
All double negation and de Morgan laws valid				

In the last chapter, we have seen that (and why) the Law of Excluded Middle fails in $N3$. $N3_f$, on the other hand, validates Excluded Middle. This means that $A \vee -A$ will never be falsified. If you check the clauses, you will find that a falsification of $A \vee -A$ would consist in a verification of A and a falsification of A, something we decided to forbid. For exactly the same reason, we see that $-(A \wedge -A)$ is valid in $N3_f$.

However, unlike $N3$, the falsificationistic $N3_f$ is paraconsistent, i.e., does not allow for the inference from a contradiction to an arbitrary statement. In terms of the Kripke semantics, if A has no value, neither does $-A$, so that it is trivial to construct a counterexample to $A \wedge -A \vDash_{N3_f} B$ (and also to $-(A \vee -A) \vDash_{N3_f} B$). A single world w in which A is undecidable and B falsified will do. I'll be coming back to the failure of Explosion once again in the next section, for it still merits further comment.

[2] Cf. last chapter's footnote 13 on the intersubstitutivity of A and $--A$.

In general, this logic can be seen as much an extension of LP as N_3 was an extension of K_3. Remembering that all classical tautologies are tautologies in LP as well, we might wonder if this holds in $N3_f$ as well. The answer is no, as $\nvDash_{N3f}$ $-((A \vee -A) \supset (B \wedge -B))$ shows (a counter model to this has one world in which no atoms are verified or falsified).

We do get contraposition in $N3_f$, but a failure of the dual $-(-B \supset -A) \vDash -(A \supset B)$.[3] This inference is probably as simple contraposition when we think about natural language conditionals. We tend to lose track of our intuitions quickly when conditionals are embedded in more complex structures.[4] Should we be unhappy to do without "If it is not the case that if not B then not A, then that implies that it is not the case that if A then B? "On the other hand, it is not easy to find perspicuous counter examples, either. If we are keen to keep it, $N_{\mathrm{AND}f}$ is an obvious alternative.

Lastly, not only modus tollens but even modus ponens fails us in $N3_f$. This, needless to say, seems to be a very dissatisfying feature. We should not, however, start to meddle lightheartedly with the falsification condition of the conditional, as we did with its verification condition in the last chapter.[5] Unlike the verification condition, our clear grasp of the falsification condition was always one of the main motivations for the introduction of falsifications.[6] And even if we could twist the conditional according to our will, there would still be the failure of the disjunctive

---

[3] It might be instructive to work out the simple counter model to this inference; it gives a good idea of where the trick of the unfamiliar consequence relation lies.

[4] Another inference that holds in $N3_f$ is Peirce's law: $(((A \supset B) \supset A) \supset A)$. This is very uncommon for a constructive logic, because its addition to positive (negation-less) intuitionistic logic leads to positive classical logic. In this case, this does not happen because modus ponens fails. However, the interest and importance of this inference lies wholly in the theoretical realm. A natural language assertion that would be best analyzed by this structure is very unlikely to be uttered by an actual speaker in a conversation.

[5] Even though I think this is true, I did play around with different ideas a bit; I won't bore you with my failed attempts to patch things up. For the curious, I'll report on the nearest miss.

The idea is to count $A \supset B$ as falsified only if $A \supset \neg B$ is verified. In terms of the relational semantics, the truth and falsity conditions of the conditional then become:

$w \Vdash_1 A \supset B$ iff for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$

$w \Vdash_0 A \supset B$ iff for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_0 B$

H. Wansing has recently been proposing these clauses in connection with various logics. His motivation for this is that they validate two principles that supposedly go back to Aristotle and Boethius. In any case, the two principles are named in their honor:

Aristotle $\neg(A \supset \neg A)$ and $\neg(\neg A \supset A)$
Boethius $(A \supset B) \supset \neg(A \supset \neg B)$ and $(A \supset \neg B) \supset \neg(A \supset B)$

These two (or rather four) principles surely seem plausible, but very few modern logics contain them as theorems. In classical logic, for example, all of them are false if A is false. Logics in which the Aristotle and Boethius principles are valid are called *connexive logics*.
While this clause looked more or less promising for my project, it has recently been found to trivialize N_3 (cf. Wansing 2010).

[6] Cf. Sect. 6.9.

syllogism ($\neg A, A \vee B \models B$) to explain away. Surely, the falsification condition of the disjunction would be even less appealing to give up.

Better, then, to find a uniform way to cope with these failures that leaves our semantic clauses intact. To that end I will introduce the idea of *incoherence* in the next section.

9.3 Incoherence

The story of incoherence, though, begins at another point that seems to have little to do with the failure of modus ponens and disjunctive syllogism. We will backtrack and take yet another look at the failure of Explosion in $N3_f$.

I argued in Sect. 7.5.1 that paraconsistency is a desirable feature in a falsificationistic logic, and I'm still sticking to that claim. When two people are arguing, and both of their assertions will be correct if the audience cannot falsify them, then we will expect to see situations in which one party asserts A and the other $-A$, and where both of these assertions will be correct. Thus the need for paraconsistency: From unfalsified premises A and $\sim A$ it does not follow that any B is unfalsified.

Something that *does* follow from such premises, though, is that $(A \wedge \sim A)$ is unfalsified. As we are examining the possibility of equating unfalsifiability and assertibility, we seem to be heading toward allowing people to contradict *themselves*, not just their opponents. Should we really be happy to allow people to utter contradictions like that?[7]

Presumably, we should like to ask the participants of our debates to maintain a certain level of coherence in their claims, and part of such coherence is not to utter outright contradictions of the form "A and not A." I'll give a formal characterization of coherence below; for now, intuitively it is approximately supposed to mean that one is not disagreeing with oneself.

Let us see some examples of what I take to be coherent and what incoherent. In a falsificationistic theory, we want to allow the following situation to pass without criticism:

A undecidable.
Alan says: "A"
Bertha says: "Not A"

[7] "Of course," the dialetheist might tell us, "because some of the contradictions are true, and what is true is surely to be asserted!" But a normal state of information will leave many more contradictions assertible than even the most liberal dialetheist would want to endorse.

What we are unhappy about is this situation:

A undecidable.
Alan says: "*A* and not *A*"

In the last example, Alan was not meeting the coherence constraint, and thus, we may criticize his assertions.

However, the charge of incoherence is different from the kind of criticism that Dummett was talking about when he spoke of falsifiability. Alan did not offend by saying something that we can falsify, because we are in no position to ascertain which of his statements, *A* or *not A*, is false. Incoherence and falsifiability are two different issues, and there are good reasons to keep them apart.

First of all, some may doubt that this kind of incoherence should at all times give rise to criticism in the first place, for example, the dialetheists (those who believe in true contradictions) among us. Second, even if we reject dialetheism, the consequences that the two kinds of flaws have might turn out to be quite different. For example, consider once again a court case. Intentionally saying something that is falsifiable under oath clearly allows the judge to convict the witness of the crime of perjury. On the other hand, simply pointing out that the testimony was incoherent while being unable to prove which part of the testimony was actually false might not be enough to give out that verdict (though if it is clear that the incoherence is intentional, and the defendant sticks to all of the incoherent claims even after prompted, he or she might get into trouble for ridiculing the court).

Even disregarding the question what constitutes a proof of perjury, the difference between incoherence and falsifiability will affect the further course of the trial. A sly lawyer can shed doubt on a witness's testimony by leading her on to contradict herself. But that does not mean that anything in particular she said will from then on count as conclusively falsified: Which part of her testimony would that be?

If we *do* want to keep apart the two issues of falsifiability and incoherence, then we should welcome the fact that the logic of falsification does not rule out incoherent statements (unless, of course, they are not only incoherent, but also *falsifiable* as well). Of course, that does not mean that we are not interested in the question whether a set of statements is logically incoherent and thus open for the second kind of criticism.

Before we go on to see what we can do against incoherence, we have to say something about the failure of *modus ponens* as well. This problem is actually analogous to the one involving the unfalsifiability of "*A* and not *A*." In the following illustration, let us assume that *B* is falsified, while nothing is known about *A*.

First of all, this situation should be unobjectionable:

A undecidable, B falsifiable.
Alan says: "*A*"
Bertha says: "If *A*, then *B*."

Both Alan and Bertha make assertions that are unfalsifiable and thus correct. Given the falsifiability of *B*, we can see that they are in thinly disguised disagreement with each other, but as we make room for faultless disagreement, this does not speak against their being correct.

However, we get into trouble if we speak as Alan does in the next example:

A undecidable, B falsifiable.
Alan says: "*A*. If *A*, then *B*."

He really should not be saying both of these things together in view of the fact that *B* is falsified. We may point out to him that he should refrain from one or the other, but we cannot put our finger on which of the two has to go, as we do not have enough information concerning *A*. In other words, everything he said is unfalsifiable, but in total, his assertions are incoherent and therefore objectionable. The only difference to the asserted contradiction is that what Alan says is not incoherent unconditionally, but incoherent relative to all states of information in which *B* is falsified.

As indicated, disjunctive syllogism gives rise to the very same phenomenon:

A undecidable, B falsifiable.
Alan says: "Not *A*"
Bertha says: "*A* or *B*."

presents no problem, even if *B* is clearly false.

On the other hand,

A undecidable, B falsifiable.
Alan says: "Not *A*. *A* or *B*."

is a problem, but it is a problem of incoherence rather than a problem of falsifiability.[8]

Again, these cases do not speak against the logic N_{3f}, but they point to a missing link in the theory. The underlying contention is that it is worthwhile to keep track of why we are criticizing other's assertions. Our logic tells us what comes out as unfalsifiable given some specific unfalsifiable items, but it tells us nothing about the coherence constraint we wish to further impose. We will have to find some independent way of accounting for coherence.

9.3.1 Getting Rid of Incoherence

So, what can be done against incoherence? Here is a proposal.[9] I'll be taking advantage of a modal extension of Nelson logic introduced in Wansing (1998), p. 149.[10]

The language is extended by a modal operator M. The statement MA is to be read as "it is consistent with what we know at present to assume A." We get the following semantic clauses for this operator:

$w \Vdash_1 MA$ iff there is an $x \geq w$, $x \Vdash_1 A$

$w \Vdash_0 MA$ iff for all $x \geq w$, $x \Vdash_0 A$

Even though we still postulate persistence of both 0 and 1 for atomic statements in these models, it is easy to come up with models that show that this does not hold for arbitrary formulas any more. However, for formulas not containing the modal operator, the heredities still hold true.

Now, how might extending our language to one with a modal consistency operator help with our problem of incoherent speech?

---

[8] The following pair, illustrating the case of modus tollens, looks completely analogous as well:

A verifiable, B undecidable.
Alan says: "Not B"
Bertha says: "If A, then B."

is fine,

A verifiable, B undecidable.
Alan says: "Not B. If A, then B."

might be incoherent. Whether it is will presumably depend on our decision about contraposition and modus tollens at Stage II.

[9] It is reminiscent of and indeed inspired by Jaśkowski's discussive logic (Jaśkowski 1969).

[10] Wansing's aim here has nothing what so ever to do with incoherence; rather, it is to supply a base for a system of non-monotonic reasoning.

The most straightforward proposal is this: We decide whether a set of statements $A_1, \ldots, A_n$ can be coherently asserted by checking whether their conjunction is possible, that is whether our current state of information supports $M(A_1 \wedge \cdots \wedge A_n)$.

Simple Modal Coherence A set of statements $A_1, \ldots, A_n$ is coherently assertible at a world w iff $w \Vdash_1 M(A_1 \wedge \cdots \wedge A_n)$

This solution indeed does away with assertions of contradictions (generally incoherent assertions), as $w \nVdash_1 M(A \wedge -A)$ for all worlds w. Assertions that are only incoherent in view of the information at a given world can also be discounted. It can be easily checked that if we have $w \Vdash_0 B$, then $w \nVdash_1 M(A \wedge (A \supset B))$.

However that might be, the modal strategy throws out too much unobjectionable material along with the things we indeed want to get rid of. From all we have seen up to now by way of motivational arguments for unfalsifiability as an assertion norm, there is no reason at all to discount assertions that are unfalsified at a world, even if we have conclusive proof that no verification for the asserted statement will ever be obtained. Let us consider the following statement, uttered in our times:

Julius Caesar had an even number of hairs on his head when he was stabbed to death.

It is unlikely to the point of certainty that we should in the future find historical records that deal with this issue or devise a method that could deliver good evidence for or against this statement. So, we may consider it verified that we will never verify the statement above. The Simple Modal Coherence criterion will thus rule this statement out as incoherent. Nonetheless, and importantly differently from the case of an asserted contradiction, this should not give us grounds to dismiss the statement.

However, what the modal statement we want to use as a filter says is that a statement is coherent iff it is possible that it will be verified at some point. This is false in the described scenario. Thus, the claim will be stigmatized as incoherent under the strictures of Simple Modal Coherence, something we should not want.

Here is a way to get around this problem.[11] It involves adding yet more modal vocabulary to our language to make reference to past states of investigation possible. Then, we will be able to say that a (set of) statement(s) is coherent if, at some point in the past, there had been a possibility that it would become verified later on. That is, we take the claim about Caesar's hair to be coherent because we can discern a point in the past, say the morning of his assassination, where the verification of it was a possibility (even if it would have been unlikely even back then that someone should have actually counted the number of hairs on his skull).

The addition to our vocabulary is straightforward: A modal operator P is defined as follows:

$w \Vdash_1 PA$ iff there is an $x \leq w, x \Vdash_1 A$

$w \Vdash_0 PA$ iff for all $x \leq w, x \Vdash_0 A$

Now, we can spell out the proposal for coherence:

[11] Thanks to Graham Priest for suggesting this fix.

Elaborate Modal Coherence A set of statements $A_1, \ldots, A_n$ is coherently assertible at a world w iff $w \Vdash_1 PM(A_1 \wedge \cdots \wedge A_n)$

One could proceed from here by trying to cram the notion of coherence into the definition of logical consequence to get a "coherent logical consequence" relation. However, the result is neither elegant nor very useful. Better, it seems to me, to keep the two tools separated. Given a number of unfalsifiable statements, $N3_f$ will tell us which other statements will be unfalsifiable. From those, together with background information on what a given speaker has already asserted, one can distill sets of statements that are coherently assertible for that speaker.

9.4 Safe Assertibility Reconsidered

Let me now fulfill a promise that I made long ago (in Sect. 7.3.1) by discussing the notion of safe assertibility in the light of the things we have learned in the meantime. It was Rumfitt's idea that a statement should be assertible iff it could be recognized that it would never be falsified, something he marked by calling such a statement "safely assertible." In adopting all the definitions that Dummett had proposed, he was able to show that by altering the consequence relation to make it transmit safe assertibility, one gets classical logic out of this setup.

I have said before that this account of assertibility seems to me to be inferior to the old notion of *pro tempore* unfalsifiability (though not necessarily as an account of falsificationistic truth). Nonetheless, I will take the idea on board in this section and see where it leads us when we develop it along the same lines as I did develop the idea of logic transmitting *pro tempore* unfalsifiability.

As Rumfitt offers no new account of the falsification conditions of the connectives, the criticisms I raised in Chap. 7 will hit Rumfitt as hard as Dummett.

My conclusion was that verification conditions are needed to make the account plausible. The notion of safe assertibility adds nothing that would make the definitions of the logical constants in N_3 less plausible, so I think we should stick to these. In order to get to a plausible logic of safe assertibility, I propose to redefine consequence thus:

$\Gamma \vDash A$ iff in every N_3 model and every $w \in W$, if $x \nVdash_0 B$ for every $B \in \Gamma$ and every $x \geq w$, then $x \nVdash_0 A$.

One may doubt whether this modification would lead to classical logic as well, and one would be right to doubt. In fact, this redefinition of logical consequence leads to a logic that coincides with none of the logics we have seen so far.

First, we may quickly come to see that all valid inferences of $N3_f$ still hold on this reformulation. Suppose that at a world, all premises in Γ are unfalsified in all subsequent worlds, and suppose that $\Gamma \vDash_{N3_f} A$. Then, at all subsequent worlds, A will be unfalsified, as the premisses are unfalsified at those worlds and $N3_f$ consequence transmits unfalsifiability at a world.

However, we will also get some new inferences, such as the unintuitive looking $-A, -(A \supset B) \vDash C$. For $-A$ to remain unfalsified means that A will never be verified and that is enough to verify $A \supset B$, thus to falsify $-(A \supset B)$; the premises cannot be both forever unfalsified.

It is unlikely that such an inference will be a welcome feature of a logic of safe assertibility. However, if one is tempted to investigate this logic further, one can easily do so by noting that the condition of eternal unfalsifiability of A is expressible in the object language of $N3$ as $\sim -A$, where $\sim$ is intuitionistic negation, defined as $\sim B =_{\text{def}} B \supset -B$. Thus, iff $\sim -A_1, \ldots, \sim -A_n \vDash_{N3} \sim -B$ holds, $A_1, \ldots, A_n \vDash B$ holds in the logic of safe assertibility.

9.5 Chapter Summary

In this chapter, I introduced the logic $N3_f$, in my view the most suitable logic for an expanded falsificationism. We were able to draw on many of the insights gathered in earlier chapters, so that the characterization went quite swiftly.

However, to account for the valid and especially the invalid inferences in this logic, some concentrated effort was needed. In particular, the failure of modus ponens and the disjunctive syllogism seemed worrying.

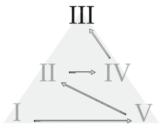
But also the paraconsistency of $N3_f$, hitherto lauded as a great feature, proved under closer inspection to harbor a certain implausibility. It seems that a falsificationistic theory has to acknowledge the unfalsifiability and therefore the assertibility of outright contradictions.

To answer all of these worries, I introduced the new notion of incoherence. Intuitively, one is being incoherent if one is disagreeing with oneself. I made this notion formally precise and showed how it applied not only to the case of outright contradictions, but also to the invalidity of modus ponens and the disjunctive syllogism.

Lastly, I went back to Rumfitt's notion of safe assertibility and showed how it, too, can be captured in the general framework I proposed.

This ends our tour through the theories that were based on only one notion in the assertoric content. What is left is to rise to the top of the pyramid, Stage III. In the next chapter, we will see whether and how verifications and falsifications may live together not only in the ingredient sense, but also in the assertoric content.

Chapter 10
Stage Three: Hybrid Strategies



10.1 Chapter Overview

We have finally reached the top of the pyramid. Here, at Stage III, we will have full interaction between verifications and falsifications in both the ingredient sense and the assertoric content. I call a strategy that allows us to combine the two notions at the assertoric level a *hybrid strategy*.

I'll be employing the findings of the last chapters, especially the last two on the expanded level. In effect, much of this chapter presents ways in which expanded verificationism and expanded falsificationism can be managed to work alongside each other. In terms of logical systems, this means that I'll be relying on the Nelson models and the two ways in which consequence can be defined on them (verifiability preservation and unfalsifiability preservation).[1]

The three main types of hybrid strategies for the combination of verifications and falsifications in the assertoric content are, as I have already announced before: The *discourse separation strategy*, the *correctness as verifiability and incorrectness as falsifiability* (*CV&IF*) *strategy*, and the *burden of proof distribution strategy*.

Discourse Separation

The discourse separation strategy simply states that both verificationism and falsificationism might have their place in our theory of meaning, albeit in different areas of discourse. There is actually not much more I wish to add to what I have said in the course of this book about this strategy. As I said before, I'm not really sure that we will be able to decide of every statement which area of discourse it should belong to. If we are making discriminations between, among others, mathematical and taste

[1] I'll speak of N_3 and N_{3f}, not meaning to imply that N_{AND} and $N_{\text{AND}f}$ are off the table. It is just that conditionals will not be an issue of importance in this chapter.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_10,

discourse, will the statement "The number of tasty things in the world is a perfect number[2]" belong to the former or the latter?

But *if* we can actually pull off a discourse separation in a clean enough way, then the issue is relatively straightforward. The task will be to inspect each area of discourse and decide which norm of assertion, be it verificationistic, falsificationistic, or possibly one of the hybrids below, fits it best. Again, in practice, this might well turn out to be very hard, but the basic idea is simple.

Correctness as Verifiability and Incorrecteness as Falsifiability

This is the idea that an assertion is correct iff it is verifiable, and incorrect iff it is falsifiable. As you will remember, Dummett explicitly rejected this strategy,[3] so we will have to address his worries. In fact, I'll start this chapter with a restatement of his arguments. Recall that he was worried about (a) statements that were neither correct nor incorrect, and (b) about truth value gaps. Surely, the two worries are related, but the relation is complicated by the many ways in which truth, falsity, verifiability, falsifiability, correctness, and incorrectness might be correlated.

The point of discussing these matters again in this chapter is not just to bring them back to memory: Given that we have seen concrete semantical proposals for Stage II and Stage IV in the last chapters, we now have a much clearer focus on some of these issues, and the appraisal of Dummett's ideas will be much more feasible.

Conversely, this discussion will supply some more philosophical insights about the logics of the last two chapters.

> Important Terminological Note:
>
> Up to now, I've been accepting Dummett's claim that there are no gaps between correctness and incorrectness. That is, what I called verificationism is not only the idea that an assertion is correct iff verifiable, but that an assertion is correct iff verifiable *and incorrect otherwise*. Likewise, falsificationism holds that an assertion is incorrect iff falsifiable *and correct otherwise*. In other words, CV&IF is neither a form of verificationism nor a form of falsificationism, because the equivalence of an assertion being correct and its not being incorrect (and of its being incorrect and its not being correct) breaks down.

Burden of Proof Distribution

Finally, I take inspiration from legal discourse and explain how verificationistic and falsificationistic ideas can be combined in the *same* area of discourse. The basic idea will be that for every assertion, the speaker either bears the *burden of proof*, or the audience has to bear it. Depending on where this burden lies, the assertion will be correct iff verifiable or correct iff unfalsifiable.

---

[2] A perfect number is a number that is equal to the sum of its proper divisors, such as 6.

[3] Cf. Sect. 5.4.

In court, it is the duty of the judge to keep track of the burden of proof. I'll try to show how it might be possible to allocate it in every day conversations.

10.2 Dummett on Hybrid Strategies

As we know from Part II, Dummett was suspicious of any attempt to construct a hybrid theory of meaning. He believed that this would inevitably lead to situations in which an assertion can come to be understood to be neither correct nor incorrect, in his view an unacceptable result.

10.2.1 Gappy Semantics

Before I turn to this worry, I should remind you of an apparently closely related one: As I mentioned in Sect. 4.6, he was wary of truth value gaps, understood as statements that are neither true nor false.

The reason this second worry should be dealt with before any attempt at a hybrid strategy is made is this: It seems to question choices I made already on the last level of the pyramid. In working with Nelson models, I accept a semantics that ostensively has gaps between its semantic values: At a typical world in a typical Nelson model, there will be statements that are neither verifiable nor falsifiable.

Luckily, it is comparatively easy to bring the gaps between verifiability and falsifiability in harmony with Dummett's view. In Sect. 4.6, we saw that Dummett saw no problem in distinguishing different *ways* in which a statement could be true or false. He allowed that these different kinds of truth and falsity, formally marked by different designated and undesignated values, would have different effects on the behavior of statements in more complex statements. That is, to distinguish between the different values is essential when it comes to the *ingredient sense* of the statement.

Dummett was talking of truth and falsity there, but we may presume that he similarly would not have had problems in discriminating different *ways* in which a statement could be correct or incorrect. And in fact, this is just what we have been doing. Up to now, the gaps in our semantics have been filed either under the label "incorrect" (in verificationistic theories) or the label "correct" (in falsificationistic theories).

10.2.2 Dummett Against In/Correctness Gaps

It would seem that our gappy semantics should not distress Dummett, then. However, there is no doubt that he clearly opposed gaps between correctness and incorrectness.

All the semantic values we choose to distinguish must either stand for a way of being correct or for a way of being incorrect.

We also saw that he thought that only one of the notions, either correctness or incorrectness, could be essential to our grasping of the *assertoric sense* of a statement.[4] The other notion would simply have to be the absence of the substantive one.

The idea that the substantive notion was actually not correctness, but *incorrectness*, was what led us to consider the falsificationistic theories of meaning in general. The argument for the primacy of incorrectness was based on Dummett's claim that finding an assertion to have been made correctly has no systematic effect, while finding it incorrect incurs a stable effect that is systematic enough to base a theory of understanding on: An incorrect assertion has to be retracted by the speaker (at least if challenged to do so by the audience).[5]

Given all this, the following list summarizes my reconstruction of Dummett's falsificationism:

Falsificationism

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Speaker need not retract statement	Correct
Gap	Speaker need not retract statement	Correct
Falsifiable	Speaker must retract statement	Incorrect

The "effect" in the first two rows is actually not an effect at all, but merely the absence of the one systematic effet Dummett allows for.

However, we already in Chap. 5 saw that sometimes less than falsifiability was needed to make a speaker retract an assertion. For example, for a mathematical assertion, it would be enough to point out that there is no (or only an invalid) proof at the speaker's disposal. Thus, the picture for the (more traditional) verificationistic theories is this:

Verificationism

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Speaker need not retract statement	Correct
Gap	Speaker must retract statement	Incorrect
Falsifiable	Speaker must retract statement	Incorrect

In either case, though, it will be enough to have a grasp of either verification conditions (verificationism) or falsification conditions (falsificationism) to grasp the full assertoric sense. Therefore, says Dummett, there is no room for hybrid strategies

[4] The assertoric sense is what decides the adequate use of the statement on its own, not as a constituent of a larger statement.

[5] Cf. Sect. 6.2.

in which assertoric sense is determined by both notions. As we know, the only way in which he considers this could happen is the CV&IF strategy: An assertion is correct iff verifiable, and incorrect iff falsifiable, so that the resulting picture would have to look something like this:

CV&IF (plain)

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Speaker need not retract statement	Correct
Gap	?	Neither correct nor incorrect
Falsifiable	Speaker must retract statement	Incorrect

But either, Dummett seems to tell us, verificationism is right, in which case the gap has the systematic effect of forcing the speaker to retract the statement, or falsificationism is right, in which case the gap will reliably lack this effect. In either case, one of the first two tables above will better display what is going on than the third one. We need to investigate which one it will be, that is, decide whether we are verificationists or falsificationists. If we decide to be verificationists, then we will do away with falsifications in the assertoric sense, if we are falsificationists, verifications will be dismissed.

Under this reconstruction, what Dummett's rejection of correctness/incorrectness gaps comes down to is this: There is one single effect the ascertainment of a statement's semantic value may have or fail to have: The speaker can be made to take his assertion back. The distinction between statements that can have this effect and those that do not (which is the distinction between incorrect and correct statements) either coincides with the distinction between unverifiable and verifiable statements or with the distinction between falsifiable and unfalsifiable statements. In either case, there is only room for one central notion in the assertoric content.

This chapter is about how we can avoid this conclusion. One simple idea is to deny that we have to make the choice between verificationism and falsificationism once and for all, that is, for all areas of discourse, we might enter into. This is the underlying idea of the discourse separation strategy: One area will be evaluated verificationistically, another falsificationistically. In general, we will need verifications and falsifications in the assertoric sense of statements to be able to do that.

Such a discourse separation strategy would not be making the central claim of CV&IF, that an assertion is correct *iff* verifiable and incorrect *iff* falsifiable. The "only if" parts of these two biconditionals would not hold. Rather, we would have the following:

> An assertion is correct if it is verifiable, incorrect if it is falsifiable, and if it is neither verifiable nor falsifiable, then it will be correct in a falsificationistic discourse and incorrect in a verificationistic one.[6]

[6] This assumes, for simplicity, that the only options we are allowing for in a single area of discourse are verificationism and falsificationism, disregarding the options yet to come in this chapter.

This seems a sensible strategy; nonetheless, let us for now suppose that we do not employ it in order to see whether we can instead make sense of a global CV&IF strategy after all.

10.3 CV&IF: Additional Effects

We have for a long time been granting Dummett that only the retraction behavior is systematic enough to be the basis of our linguistic understanding. Even when we considered verificationistic theories, we noted that we could view them in this light, by holding that an assertion that is not verifiable needs to be retracted.

That it is unnecessary to draw finer distinctions than *verifiable/not verifiable* or *falsifiable/not falsifiable* at the level of assertoric content seems in large part to be caused by only allowing for this *one* systematic effect. If we could find a systematic effect to go with verifiability/correctness in addition to the one to go with falsifiability/incorrectness, then there would be a strong motivation for a hybrid strategy: It would be necessary to employ both verifications and falsifications in the assertoric sense. The gap in the semantics will mark the absence of both effects.[7]

As it happens, we had already seen Dummett himself suggest such an effect on page 106: If a statement turns out to be correct, then the audience has to openly endorse it as well (or at least, the speaker has an effective means of making them endorse her statement).

CV&IF (with added effect)

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	No systematic effect	Neither correct nor incorrect
Falsifiable	Speaker has to retract statement	Incorrect

Let us try this out on an example: Suppose I say "It will rain tomorrow." If there is no way of determining today whether it will rain tomorrow, there is no need for you to take my word, but neither is there a reason for me to retract my statement. But once tomorrow comes, either I'll be able to make you endorse my statement,[8] or you will be able to make me retract it.

[7] Of course, we *might* also find that assertions that are neither correct nor incorrect have some systematic effect. However, the minimal requirement for the necessity of verifications and falsifications in the assertoric content is to have two: One for correct and one for incorrect assertions.

[8] Of course adjusting for indexicality.

10.4 Verificationism and Falsificationism Again (with Additional Effects)

The example and the account of it I just gave actually does not sound all bad. However, the idea that this could be a global account, not one for just *some* areas of discourse, is not very plausible.

This is because we have already seen some cases in which a statement can be incorrect without having to be falsifiable. Take, once again, the mathematician who claimed that Goldbach's conjecture is true, but had no proof (or a faulty proof) of it. The hearers may request that he take back that assertion without themselves being able to falsify it (i.e., without being able to provide a counter example).

On the other hand, it seems that if a statement is falsified, then the speaker has *not only* to withdraw it. When prompted, he even has to assent to the negation of the statement he uttered.[9] The mathematician in the example above is under no obligation to assent to the negation of Goldbach's conjecture.

One might think that this has nothing to do with the assertoric content of the statement, but rather with the ingredient sense: Where the statement is falsified, the negation is assertible. But here we are saying more: The speaker's obligation to endorse the negation of what he said is a consequence of the very same challenge on his assertion that makes him retract the statement.

It seems then that in this, the paradigmatic verificationistic case, there are two types of being incorrect. The first is to utter something unverifiable, which entails the obligation to retract. The second is to utter something falsifiable, which incurs the additional obligation to assert the negation of the statement.[10]

These three cases, correctness and the two types of incorrectness, map neatly onto the values we had seen in our semantics in the last two chapters. A proposition that receives value 1 is verifiable and thus correctly assertible; a truth value gap means that an assertion would be incorrect and should have to be retracted if this is pointed out. If the value is 0, the statement is falsified and the speaker should not only retract his assertion, he should assert the negation of his statement.

---

[9] As we are building on logics that have toggle negations, we also have this plausible outcome: If the asserted statement was a negated one, $-A$, the speaker has not only to endorse its negation $--A$ but also the unnegated statement A if his assertion is falsified.

[10] At this point, we may again ask whether it really is the right notion of "verifiable" we are working with. We are, ever since Sect. 3.8.4, assuming that a statement is verifiable iff it is decidable (that is, we have a decision method for it) and, were we to carry out this method, we would find the statement verified.

If a statement is verifiable in this sense, and we have not yet carried out the decision method, should we want to say that the speaker can make the audience accept his statement? I think this is correct, for the way to do it simply is to carry out the decision method and displaying the result. The audience's endorsement is not a direct consequence of his making a correct statement, but he can effect it if he wants to.

Verificationism (with added effects)

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	Speaker must retract, but need not assent to negation	Incorrect
Falsifiable	Speaker has to retract statement and concede negation	Incorrect

Again, this is of course presupposing the verificationistic dictum that an assertion is correct iff it is verifiable and incorrect otherwise.

But we can also supply a similar mapping for the falsificationistic theory: Again, a verifiable statement is such that the speaker can make the hearers assent to his assertion. A falsification, likewise, means that the speaker has to assent to the negation of his statement. So far, everything is just as in the verificationistic case. But the gap, of course, means that the asserted statement is, at least for now, correct. However, just as there was a difference between the two kinds of incorrectness in the verificationistic case, here we see a difference in the two ways a statement can be correct. If such a statement is correct only because it cannot be falsified, and not because it can be verified, then the speaker need not retract it. However, the further consequence we see in the case where the statement is verified, viz. that the audience has to endorse the statement in question, is absent here.

I might say that spinach is tasty and be correct in saying so, but you still would not have to agree that spinach is tasty. In fact, you might correctly assert that spinach is not tasty. So, if we entertain falsificationism as a possible norm of assertion, it seems that not in all cases, a correct assertion will have the positive consequence we are considering, viz. that the audience will have to endorse the asserted statement. It is only a consequence of those correct assertions that are correct because the statement is verified.[11]

Here, then, we have the following trichotomy in the case of a falsificationistic norm of assertion:

Falsificationism (with added effects)

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	Speaker need not retract, but hearers do not have to endorse statement	Correct
Falsifiable	Speaker has to retract statement and concede negation	Incorrect

10.5 A Tripartite Setup

Now, let us suppose that we want to use a discourse separation strategy that explains some areas of discourse verificationistically (such as mathematical talk), some

[11] If what I have said earlier about taste statements and about expanded falsificationistic meaning theories is right, then "Bile is not tasty" would be such a verified statement.

falsificationistically (such as taste talk) and all the rest via CV&IF. In the last list for the CV&IF strategy on page 180, we had not yet filled in the stronger negative effect of a falsification:

CV&IF (w/ added strong negative effect)

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	No systematic effect	Neither correct nor incorrect
Falsifiable	Speaker has to retract statement and concede negation	Incorrect

Now, one might think that this is actually just a terminological variation of the last list, the one for falsificationism. One of the cases we chose to call a *correct* statement in the falsificationistic theory is now called *neither correct nor incorrect*. Whereas I said there that the speaker need not retract his statement and the hearers do not have to accept the statement, I here write that the statement has no systematic effect, which seems to come to the same thing.

In fact, though, it *does not* come to the same thing. The crucial difference lies in the effect that has been our main focus all along: The obligation to retract.

It is a scopal difference between "systematically no effect" and "no systematic effect." While falsificationism tells us that an undecidable statement *lacks, quite systematically, the effect* that the speaker has to withdraw the assertion, in the CV&IF cases, we merely find that the obligation to retract is *not a systematic effect*. Those assertions that would come out neither correct nor incorrect *might* have no effect, but they *might* also have the effect that the speaker has to retract the statement. This will depend on the context of the assertion. Consider the following example:

While on holiday, we drive across the Golden Gate bridge. I say to you: "This bridge will not break down in a hundred years." I have no evidence for this that would add up to a verification, but you have no counter evidence either, and my assertion is allowed to stand.

Now consider Mr. Johnson, an engineer hired by the city of San Francisco to check the safety of the bridge. After some time and having received a large sum of money, Mr. Johnson declares: "This bridge will not break down in a hundred years." If it turns out that he has not enough evidence to back up this claim, that he maybe even did not make any measurements and calculations, his employers will certainly be able to make him take back his assertion.

I think that it would be very implausible to claim that what we are dealing with here are two distinct areas of discourse. The difference, rather, seems to lie in features of the context, in this case *who* made the assertion in *which* kind of circumstances.

So, a more explicit list for CV&IF to use alongside the verificationistic and the falsificationistic one would look like this:

CV&IF (elaborate version)

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	Hearers do not have to endorse statement Speaker does not have to endorse negation Speaker might have to retract statement (depending on context)	Neither correct nor incorrect
Falsifiable	Speaker has to concede the negation	Incorrect

I take the resulting tripartite picture to be mildly satisfying. However, it would be *much more* satisfying if we could get a grip on which features of context decide how gaps are to be evaluated and describe their effects in a systematic way.

10.6 Unifying the Account

Indeed, if we could do this, we could unify our account of verificationistic, falsificationistic, and other areas of discourse. Note that the Golden Gate examples could have been described in the following way: The first assertion is judged according to the verificationistic norm, the second one according to the falsificationistic one. The only difference to the earlier description is this: Above, we called both assertions *neither correct nor incorrect*, now the first one would come out *correct* and the second one *incorrect*. The predicted effects, however, are exactly the same: If the statement is undecidable, then in the first example, the hearer does not have to accept the statement, and the speaker does not have to retract it. In the second example, the speaker can be made to retract the assertion, but does not have to endorse the negation.

Seen from this alternative perspective, what we are hoping for is this: A method of establishing whether a given assertion should be judged according to the verificationistic or the falsificationistic standard, given the context and the type of utterance. Sometimes, it will be enough to know what the assertion is about (mathematics, taste talk), sometimes more information will be needed.

The result would be this:

Unified account

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	Hearers do not have to endorse statement	Correct or incorrect (depending on context)
	Speaker does not have to endorse negation	
	Speaker might have to retract statement, *depending on context in a systematic way*	
Falsifiable	Speaker has to concede the negation	Incorrect

I'll try to get a grip on this picture in the next section. The conceptual tool that I'll employ is the *burden of proof*. Whether an undecidable assertion has to be retracted depends on whether or not the speaker has to bear the burden of proof. As I mentioned quite a few times[12], the blueprint for the management of the distribution of the burden of proof is legal discourse.

10.7 Burden of Proof Distribution

Occasionally, there is talk of "the burden of proof" in all kinds of discussions and disputes, but nowhere is the concept more explicitly dealt with than in legal theory. To know where the burden of proof (BoP) lies is essential to adjudicate between two opposing parties. In countries where a lay jury will give the final verdict, the explanation of the concept of the BoP is an important part of the instructions the jurors receive before the trial.

In criminal cases, the situation is especially clear. From the outset, the defendant enjoys the *presumption of innocence*: the prosecution has to prove his guilt; if they fail to do so, the defendant will go free. Both his testimony and his plea (we assume he pleads "Not guilty") will stand as long as what he says is not disproven.

The presumption of innocence represents an assessment of the risks of an erroneous verdict. A society that would rather let a guilty man go free than to convict an innocent one will have the presumption of innocence built into its legislature in some form.

As L. Laudan writes, it follows almost automatically from the presumption of innocence that the prosecution has to bear the BoP:

> If the state bears the full burden of proof, then, of course, one might say, the defendant is presumed innocent. Contrariwise, if the defendant is genuinely presumed innocent, then it naturally follows that the state must defeat that presumption by proving his guilt (Laudan 2006, p. 90).

However, the BoP does not necessarily have to stay on the prosecution. The course of a trial might shift the BoP away from them and onto the defense. For example, if

[12] E.g., in Sect. 6.6.

the defendant pleads *not guilty on grounds of insanity*, then, depending on country and state, he might have to bear the burden of proof for this claim.[13] It is the task of the judge to keep track of, decide on, and inform the jury about the BoP.

I submit that, in view of what we have seen in the earlier parts of this work, it is quite natural to say this about the statements defense and prosecution make: If a speaker has to bear the BoP, then his assertions will be judged according to the verificationistic standard; but if the burden is on the other party, then the assertions will be judged according to the falsificationistic standard.

10.7.1 Taking the BoP Outside of Legal Discourse

This might make for a promising start on a model of legal discourse. We might employ the discourse separation strategy at this point and claim a small victory. But of course the real aim is to step beyond the boundary of legal discourse and claim that similar mechanisms are at work outside court as well.

The adaptation is not perfectly straightforward, however. As we have seen, there are clear rules for the distribution of the burden of proof at court. There are default distributions and rules for when a shift occurs. Most importantly, there is always an arbiter present who will enforce these rules or make decisions about the BoP in cases where the rules give out. This arbiter is the judge.

Now, problematically, such an arbiter is missing in most situations outside of court. This is actually one of the reasons to take an issue to a civil court. For example, at the heart of many libel cases is the question whether a speaker had to bear the BoP when making certain assertions.[14]

---

[13] In many states of the USA, this shift of the BoP was legislated after a famous case in 1982. John Hinckley, a young man who was desperate to make a favorable impression on actress Jody Foster, tried to do so by shooting president Reagan and members of his security staff. His plea of not guilty on grounds of insanity was successful, not because he could prove his insanity, but because the prosecution could not prove his *sanity*. After this unpopular outcome, many states changed their laws so that a man in similar circumstances would be convicted.

[14] Currently, there is a controversy about the British libel laws that focuses exactly on the issue of the distribution of the burden of proof. The case that put this law into the limelight is between well-known science author S. Singh and the British Chiropractic Association. When Singh questioned in an article the legitimacy of the claims made by members of the association that chiropracticians could heal all sorts of heavy illnesses that had ostensively nothing to do with the parts of the body they treated, they sued him for libel.

Publicity-wise, this must have been the worst decision the BCA ever made, but in purely legal terms, they did have a point. The British libel law, unlike libel laws in other countries, could well be interpreted to say that the BoP was on Singh. This meant that he either had to retract what he said in an newspaper article about the issue, or to *prove that chiropractics could not heal those illnesses*.

In fact, this was the verdict in the first trial on the matter. Singh did not back down, took the case to a court of appeals and won in the end. The case attracted so much attention that it might well lead to a revision of the controversial law.

See: http://www.guardian.co.uk/science/2010/apr/15/simon-singh-libel-case-dropped, last retrieved on May 21st 2013. Also on May 21st 2013, Singh's homepage http://www.simonsingh.net/

However, normally conversations do not go to court, and the question which party has to bear the BoP needs to be resolved without professional aid.

What can we say about BoP distributions in normal conversations? To suppose that the burden of proof is flexible in the sense that it may at some time lie with the speaker and at others with the audience seems to entail that this feature is inherent in the structure and the context of the discourse.

There might be several possibilities this could play out theoretically. I'll sketch what I take to be the most promising idea: The burden of proof is part of what D. Lewis called the *conversational score*.[15]

10.7.2 The BoP in the Conversational Score

There are many things that need to obtain in order for a typical assertion to be unobjectionable. For example, an assertion of

He is rather tall.

will not be felicitous if there is no male in the preceding conversation or in the vicinity of the speakers who is salient enough for the hearers to work out that "he" must refer to that person. Other things that might influence the adequacy of an assertion include salient points of reference, standards of precision for vague terms and existential and other presuppositions.

Lewis now suggests that all these items are recorded in an abstract entity he calls the *conversational score*, and that the participants have to mentally "keep score" in order to be able to follow and participate in the conversation. For example, they will have to remember which salient male persons came up during the conversation, so they will be able to figure out whom "he" might refer to.

Now, I claim that just as what objects are salient in the conversation, which standards of precision apply etc., the *distribution of the burden of proof* is entered and kept track of in the conversational score.

Like any aspect of the conversational score, where the BoP lies can change during the conversation. Much of the time, such changes in the conversational score are appreciated by the speakers quite unconsciously. It is the hope of the theoretician to find clear and, at least in principle, systematizable cues to these changes.

An important phenomenon that Lewis discussed when he introduced the idea of a conversational score is *accommodation*: If an assertion is made that would be objectionable given the present conversational score, then there are two possibilities: Either the incorrectness is addressed and challenged by another participant. Failing that, the score is tacitly corrected to make the assertion felicitous.

An example makes this clear: Suppose someone said to me

(Footnote 14 continued)
collects many links to news coverage, while the BCA's homepage understandably makes no mention of the case.

[15] Lewis (1979).

By the way, have you finally stopped losing against that half-wit Jerry in your chess club?

In order for this to be an unobjectionable statement, the conversational score would have to (a) support a presupposition, namely that I was in the past usually losing to Jerry, and (b) fix the standards of applicability of the rather vague "half-wit" in such a way that Jerry falls under it.

Let us take it that neither (a) nor (b) was the case before the assertion was made. Now, unless I protest ("What do you mean, I never lost against Jerry!" or "Well, I would not call him a half-wit, after all he almost won the club championship three years ago!"), the conversational score will change so that (a) and (b) are now fulfilled. The conversation will now carry the presupposition that I used to lose to Jerry, and the standards of application of the term "half-wit" are now such that it applies to Jerry. This is the phenomenon of accommodation.

Presumably, the same phenomenon can be observed when it comes to the BoP. If I make an assertion that is clearly only correct if the burden of proof is on the audience, and the audience fails to protest, then this will mean that my assertion will stand (as long as it is not falsified). Bullying others into quietly accepting the BoP can be a formidable debating technique.

Of course, accommodation is not mandatory. There might arise a dispute about where the burden of proof lies, and it is not clear that this question can always be decided. Some of the time, it may be possible to find a neutral party that plays the role of an arbiter, just like the judge in court room debates. At other times, this might not be a viable option.

Some basic insights on how people perceive BoP distribution based on the structure of conversations have already been gathered by experimental psychologists (Bailenson and Rips (1996); Bailenson (2001)). As one would expect, many factors play into this, such as the controversiality of the claims, the conviction with which they are expressed and the mere order in which arguments are presented. But these results are not yet detailed enough to give a precise indication of how the BoP will distribute in any given conversation. In fact, fascinating as this research is, it is unclear that this is an attainable goal at all.

What, then, if both parties just will not budge and refuse to accept that the burden of proof lies with them and not the opponent? It seems that in such cases, communication might well break down at some point. Everyone involved will realize that further discussion is useless.

Now, this seems to me to be something that is in fact happening. Some disputes just go nowhere, because they are not resolvable precisely because no one accepts the burden of proof. Some people, indeed, seem to never be prepared to accept this burden.[16] The most sensible thing to do might well be to stop talking and listening to them. In some special cases, they might say something that is legally bound to the burden of proof. For example, going around and proclaiming others to be child molesters on the sole basis that these people cannot prove that they have never approached a child in an indecent manner has adverse legal consequences. But in

[16] Certain political commentators come to mind.

general, there seems little else to do against such behavior than to decline to enter the conversation.

There is obviously much more to be said about the burden of proof, but I will leave further explorations for a later occasion. Let me end this section by making a last observation about BoP distribution.

One feature that will in many discussions be different from criminal court cases is this: In court, if one party is under the burden of proof, then (a) it has to make *provable* assertions and (b) it has to *disprove* the assertions of the other party to show them incorrect. That is, the party bearing the burden bears it, whether it is in the role as speaker or the role of the hearer.

This is not necessarily the case in other discourses: At a mathematician's congress, everyone bears the burden when *uttering* mathematical statements, but not when sitting in the audience. At the annual meeting of the Conspiracy Theorist Association next door, no one might bear the burden when making claims, and everyone would bear it when listening to their peers.

A conversational score will determine whether a speaker making an assertion has to bear the burden or not. If, and only if, she does not have to bear the burden, the hearers will have to bear it, but only in their role as hearers. If they proceed to make assertions of their own, the issue must be decided anew. Given that the BoP distribution might shift during a conversation, maybe the best way to model all this would be to say that the BoP attaches to single assertions rather than to speakers.

10.7.3 Summing Up the BoP Strategy

I think that the burden of proof distribution strategy is the most interesting and appealing of the hybrid strategies. However, it also leaves us with a very difficult task, namely to describe the mechanisms in which the BoP is distributed in much more detail than I have been able to do here.

To sum up what I do want to say about it here: There is a burden of proof that either lies on an assertion or does not lie on it. This might depend on the kind of statement (mathematical, relating to taste), or it might depend on the course of the conversation it is part of. In this case, it depends on a burden of proof parameter in the conversational score. Depending on whether it does or does not lie on the assertion, the assertion will have consequences according to the verificationistic or the falsificationistic table above, which can be combined into the following single table:

Let us see how this strategy measures up against Dummett's fears and assumptions about a hybrid strategy.

Unlike what he claimed in the earlier texts, there is more than one systematic effect the ascertainment of the semantic value of a statement can have or fail to have.

Nonetheless, the BoP distribution strategy keeps the dividing line between incorrectness and correctness where Dummett drew it, between those statements that have to be retracted and those that are allowed to stand. As a consequence of this, it does

Burden of proof distribution

Semantic value	Effect on linguistic interaction	Verdict
Verifiable	Hearers have to endorse statement	Correct
Gap	Hearers do not have to endorse statement	Correct or incorrect (depending on BoP)
	Speaker does not have to endorse negation	
	Speaker has to retract assertion iff bearing the BoP	
Falsifiable	Speaker has to concede the negation	Incorrect

not allow statements that are neither correct nor incorrect. Thus, we are complying with Dummett's wish here.

However, unlike what Dummett thought a hybrid strategy would have to be, the burden of proof distribution strategy is not a CV&IF strategy: The "only if" directions of the two biconditionals break down.

10.8 Hybrid Consequence

We can also use the BoP to draw logical consequences in a logical system I'll call HYBRID CONSEQUENCE. Suppose we want to know which statements can correctly be asserted given the current distribution of the BoP and the correct assertions that have been made in the course of the conversation. These earlier correct assertions will fall into two categories: Those made under the BoP, and those made not bearing the BoP. I'll write Γ for the set of statements correctly asserted under the BoP, and Δ for the set of statements correctly asserted without the BoP. We will also have two classes of consequences we will be interested in, depending on whether the next statement we wish to make will be made under the burden of proof or not. Thus, logical consequence will relate the ordered pair of sets of statements Γ and Δ to other pairs of sets of statements, Φ and Ψ. If the burden of proof will lie on the consequence, then Φ will contain a single statement and Ψ will be the empty set, and the other way around if the burden of proof will not lie on the consequence. By "the consequence," I simply mean the statement that is in the singleton Φ or Ψ. I will write $\Gamma \mid \Delta \vDash_{\text{hyb}} \Phi \mid \Psi$ to denote HYBRID CONSEQUENCE, and simply omit any of Γ, Δ, Φ and Ψ in case they are empty.

Using the same semantical setup as in Chap. 8, i.e., the Kripke semantics for N_3, we redefine logical consequence:

> $\Gamma \mid \Delta \vDash_{\text{hyb}} \Phi \mid \Psi$, iff in every model and every $w \in W$, if $w \Vdash_1 B$ for every $B \in \Gamma$ and if $w \nVdash_0 C$ for every $C \in \Delta$, then $w \Vdash_1 D$ for every $D \in \Phi$ and $w \nVdash_0 E$ for every $E \in \Psi$.

The idea is that Γ collects all that we know is verified, because correctly uttered under the BoP, and Δ all that we know is at least not falsified, because uttered correctly by speakers not bearing the BoP. Whether or not the BoP is on the conclusion will decide whether we are interested in conclusions before or behind the bar. So, for example, the consequence A of an inference $\Gamma \mid \Delta \vDash_{\text{hyb}} A \mid$ will be such that we can utter it correctly if the BoP is on us, given that Γ and Δ collect verifiable and unfalsifiable statements, respectively.

Looking at this example, one may wonder whether the information recorded in Δ, which is essentially information about unfalsified statements, could be relevant at all if the conclusion has to be verified. Indeed, it does make a difference, as the following example will show. Let $A \vee B$ be the only element of Γ, and $-B$ the only element of Δ. Then, we will find that $\Gamma \mid \Delta \vDash_{\text{hyb}} A \mid$, while $\Gamma \mid \nvDash_{\text{hyb}} A \mid$. Intuitively speaking, the BoP-bearing speaker who correctly asserted $A \vee B$ had to have either a verification of A or a verification of B. The latter case is excluded by the fact that someone else, not bearing the BoP, was able to utter $-B$.

Now, if we know our assertion will not be made under the BoP, we will be interested in the inferences of the form $\Gamma \mid \Delta \vDash_{\text{hyb}} \mid A$. In this case, we should also check the conclusion for internal coherence using the Elaborate Modal Coherence tool I introduced in Sect. 9.3.1. However, note that the tool should *not* be employed to purge Δ of incoherent statements. Even though, e.g., an asserted contradiction is incoherent, it may give us valuable information about what has and what hasn't been falsified.

It is easy to see that $\Gamma \mid \Delta \vDash_{\text{hyb}} A \mid$ implies $\Gamma \mid \Delta \vDash_{\text{hyb}} \mid A$. It is equally obvious that $\Gamma \mid \vDash_{\text{hyb}} A \mid$ iff $\Gamma \vDash_{N3} A$ and $\mid \Delta \vDash_{\text{hyb}} \mid A$ iff $\Delta \vDash_{N3f} A$.

Inspecting the semantic clauses of the connectives shows that $\mid \Delta \vDash_{\text{hyb}} A \mid$ will hold only if A is a tautology of N_3. On the other hand, $\Gamma \vDash_{N3} A$ implies $\Gamma \mid \vDash_{\text{hyb}} \mid A$, but we also get some new inferences here, such as $A \supset B, \neg B \mid \vDash_{\text{hyb}} \mid \neg A$.

The following list collects some other inferences of interest:

$A \vee B \mid \neg A \vDash_{\text{hyb}} B \mid$	$A \vee B \mid \neg A \vDash_{\text{hyb}} \mid B$
$\neg A \mid A \vee B \nvDash_{\text{hyb}} B \mid$	$\neg A \mid A \vee B \vDash_{\text{hyb}} \mid B$
$A \mid A \supset B \nvDash_{\text{hyb}} B \mid$	$A \mid A \supset B \vDash_{\text{hyb}} \mid B$
$A \supset B \mid \nvDash_{\text{hyb}} \neg B \supset \neg A \mid$	$A \supset B \mid \vDash_{\text{hyb}} \mid \neg B \supset \neg A$
$\mid \neg(\neg B \supset \neg A) \nvDash_{\text{hyb}} \mid \neg(A \supset B)$	$\neg(\neg B \supset \neg A) \mid \vDash_{\text{hyb}} \mid \neg(A \supset B)$
$A \supset B \mid \neg B \nvDash_{\text{hyb}} \neg A \mid$	$A \supset B \mid \neg B \vDash_{\text{hyb}} \mid \neg A$
$\neg B \mid A \supset B \nvDash_{\text{hyb}} \neg A \mid$	$\neg B \mid A \supset B \vDash_{\text{hyb}} \mid \neg A$
$A \supset B \mid A \nvDash_{\text{hyb}} B \mid$	$A \supset B \mid A \nvDash_{\text{hyb}} \mid B$

Admittedly, the above has a somewhat complicated look to it. Does a relation between ordered pairs of sets of statements add up to a *logic*? More, could such a blatantly pragmatic mechanism as the distribution of the burden of proof have any claim to be at the *heart* of logic? Can the inferences this setup churns out really have anything to do with our understanding of the logical constants? Isn't logic supposed to provide a certain path form one truth to the next, rather than a complicated recipe for finding things one can get away with? By the way, what happened to truth anyway in this account?

These are good questions; however, it seems that a charge of giving an account of logic that is "too pragmatic" is not trivially easy to bring home against a philosopher whose avowed credo is "meaning is use." To use a statement competently is to be able to decide when to make it. And if this systematically involves knowing where the burden of proof lies, then the way this burden acts on the logical constants is something that has to be understood as well, with all the complexity that entails for logical consequence.

Moreover, once we move into an area of discourse where the burden of proof is fixedly on the speakers or the hearers, the account of logical consequence above becomes much easier: It will just coincide with N_3 or N_{3f}, respectively.

But what about the last of the questions above? What about truth in all this?

10.9 Whatever Happened to Truth?

Indeed, we have lost the notions of truth and falsity from view a little. We have seen that often, Dummett thought that truth and correctness coincide. However, to tie truth to correctness and falsity to incorrectness in the present account would mean to make these notions dependent, not only on the state of investigation, but also on the burden of proof.

Now, a notion of truth that fluctuates with the burden of proof truly deserves the label "anti-realistic." In fact, it comes quite close to claim that truth is for a large part a social construction: We would be "making" many of our truths and falsehoods by conducting our conversations in a certain way.

I think that at this point, many will feel that there is too much strain on our intuitive concept of truth to warrant the label. If we *must* cram truth and falsity into our picture, then I think we should simply tie truth to verifiability and falsity to falsifiability.[17]

Hybrid truth

Semantic value	Truth/falsity	Verdict
Verifiable	True	Correct
Gap	Neither true nor false	Correct or incorrect (depending on BoP)
Falsifiable	False	Incorrect

Dummett would, as we know, dislike the fact that there is a space for statements that are neither true nor false. But given that we have discerned quite different effects for verifiable, falsifiable and gappy assertions in this chapter, his argument seems not too persuasive. There seems to be no good reason any more to maintain that the effect of all types of designated (undesignated) values is the same as long as the statement is not constituent of a larger statement.

Nonetheless, there is still quite enough to upset intuitions about truth and falsity in this picture: Both notions, as we had seen already in Sect. 3.8, will be tensed: The more we find out about the world, the more will become true or false. Furthermore,

[17] And if we feel that we just have to keep truth and correctness (and falsity and incorrectness) correlated, then we should move back to the CV&IF strategy.

the disquotational scheme will fail: We want to make room for correct assertion of an undecidable *A*, even if "*A* is verifiable," and hence "*A* is true" would be unassertible even if the BoP was not on the speaker.

For what it is worth, personally, I would advise the constructivist to stay away from the notion of truth as far as possible. Any constructive attempt to give an account of truth that tries to be answerable to intuitions we may have about this concept, or to respect the pretheoretical use we make of it, is just quite unlikely to succeed.

This was also an early impulse of Dummett's: Truth is an inherently realistic notion, and a constructivist should not employ truth conditions in order to account for the meaning of statements. Correctness conditions give us all we need to explain meaning and logical consequence. For me, this is still the most promising line for the constructivist to take, even if it is far from unproblematic.

If, on top of an account of meaning and logical consequence, truth and falsity must be accounted for, then I would suggest to use an eclectic strategy, such as the ones sketched in Sect. 2.11. As I wrote there, the metaphysical harvest we might reap from such an eclectic account will presumably be meager, but metaphysical conclusions were never my main aim.

What instead I take the main results and contributions of this book to be will be summarized in the last part.

10.10 Chapter Summary

In this chapter, I have shown how verifications and falsifications can play an exactly equal part on all levels of meaning. I showed that there are more systematic effects the ascertainment of semantic values might have than Dummett allowed for, and that it is necessary to grasp both verification and falsification conditions to appreciate all these effects.

The three different ideas on how to combine verifications and falsifications into a hybrid strategy were the discourse separation strategy, the CV&IF strategy (the one Dummett considered) and the burden of proof distribution strategy. I claimed that the last strategy is the most appealing one, even if I could not supply a complete account of the shifts of the burden of proof we find in everyday discourse.

I also showed that the burden of proof can be worked into the account of logical consequence. However, only in those cases where the burden of proof is fixed do we get something that conforms to our traditional expectations of what a logical system might look like formally. I ended the chapter by suggesting the most plausible way the notions of truth and falsity could be located in the account, but I advised the constructivist to try to make do without these concepts.

Part IV
Summary

Chapter 11
Summary

In this last part, I would like to collect some of the main results of this work and make some general philosophical observations about them.

11.1 Resume

I presented the material in three large parts entitled BACKGROUND, FALSIFICATIONS and LOGICS.

Starting out on the background, I tried to give a concise but useful characterization of what the most important aspects of Dummett's constructivist program are. This included a larger overview of his wide-ranging aims, such as building a base from which it would be feasible to tackle metaphysical problems, primarily disputes between realists and anti-realists of various stripes. Though these results are fascinating, I concentrate on the basis he sets up for these considerations: The constructivist theory of language and logic.

The next chapter was devoted to the intuitionistic program, which incorporates the most important ideas Dummett builds on. I described the philosophical motivation and the early developments of intuitionistic logic and showed how the semantics of intuitionistic logic can be seen as a paradigm for Dummett's setup. However, not each and every one of the various semantics for intuitionistic logic will serve the Dummettian purpose. I argued that a Kripke semantics, supplemented by the classic BHK explanation, has some chance of being up to the task. I strove to give quite a thorough introduction to these matters, as the logical systems I introduce in the third part essentially draw on them.

A third chapter was needed to get some basics on gappy and glutty semantics clear. The concept of paraconsistency was explained and given some motivation. Importantly, I stressed that paraconsistency does not have to come bundled with dialetheism, the metaphysical doctrine that there are true contradictions. The alternative view that

A. Kapsner, *Logics and Falsifications,*
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9_11,

paraconsistency results from assertible truth value gaps, analetheism, was presented as well.

The second part contains most of the exegetical work in this book. I try to collect and make sense of the things Dummett says about the role of falsifications in a semantical theory. To get some order into matters, I discerned five different stages, which were ordered in a pyramidical diagram. This diagram was supposed to facilitate orientation in the third part and to point out the axes along which the different positions could vary.

Vertically, with increasing height, the amount of interaction between verifications and falsifications increased. At the lowest level, there was no such interaction, simply because only one central notion was allowed (either verification or falsification).

At the intermediary level, there was one central notion that was responsible for the assertoric content of a statement, i.e., for determining under which circumstances an assertion is correct. However, there was a mixture of verifications and falsifications employed to determine the contribution the sense of a statement made to the sense of a complex statement containing it.

Finally, the apex of the pyramid was where both concepts enjoyed full interaction.

The left–right axis of the pyramid marks two different answers to the question: What makes an assertion correct? On the left side, we find the *verificationistic* theories, in which the answer is this: An assertion is correct iff it is verifiable. On the right hand side, however, the answer is that an assertion is correct iff it is unfalsifiable. I call theories that give this answer *falsificationistic*.

As this defining thought of falsificationism is rather little developed in Dummett's writings, even though it takes a central place in some of his most important works, I spend the second chapter of this part on it. I quickly come to argue that falsificationism as a general theory, applicable to all areas of discourse, does not seem very plausible. However, I try to find some examples that make the proposal plausible. The most clear-cut example here is taste talk, and generally areas of discourse in which faultless disagreements are possible.

The third part of the book, entitled LOGICS, finally addresses the question which concrete effects on logical principles the introduction of falsifications brings. I start with the one logic that we find in Dummett's own exposition, dual intuitionistic logic. This logic is supposed to underwrite a pure falsificationism, where all meanings are given in terms of falsification conditions only. Although an interesting approach, ultimately I find it unsatisfying, as it cannot give a plausible account of logical vocabulary.

I proceed in the next chapter to what I call expanded verificationism, where both verifications and falsifications can be employed to determine the meaning of the logical constants. I found that the logic that most naturally flows from this setup is one of the Nelson logics, N_3 (I dismiss the glutty variation N_4 early on). I discuss some alterations on the Nelson conditional, as the failure of contraposition and modus tollens are a bit unsatisfying. However, in the end, I chose to stick with the original account.

Although I could not claim a complete knock out victory of N_3 over intuitionistic logic, I must say that the superiority of the Nelson account over the intuitionistic one is among the claims I make that I feel strongest about. I think that constructivists would do very well to pine for revision along Nelson lines rather than intuitionistic ones, both because of the theoretical reasons given above and simply because the proposed changes would, I believe, appear to the uninitiated much less bizarre and impractical.

In the next chapter, I show how the Nelson ideas can be adjusted to fit the falsificationistic requirements. Given what has been established up to that point, the characterization of the logic is straightforward. I call the logic that I thereby create $N3_f$. However, there is a problem in getting people to speak coherently. What I mean by this is that even if assertions are correct as long as they are not falsifiable, we should not have people contradicting themselves. I propose a strategy to deal with this problem, and I show how to implement it formally.

In the last chapter of the third part, I discuss the strategies that mix verifications and falsifications without restraint. I call those strategies the *hybrid strategies*. The chapter starts out by defusing certain worries Dummett had about such a strategy. Then I play through some possible approaches, and I end up liking the following account best: An assertion is judged either according to the verificationistic norm or according to the falsificationistic one. Which one of the two it will be is decided by the distribution of the *burden of proof*. Whether or not this burden lies on the speaker is a feature that is determined by the context in a systematic way. I make room in the conversational score for a burden-of-proof parameter and give some first hints on which features of a conversation might have an effect on this parameter. However, there is no doubt that there is much further study needed to vindicate this approach.

11.2 What is Constructivity?

I would like to end this work with some general philosophical musings about the material I presented. In particular, I would like to leave you with two questions. The first is this: What makes a logic constructive? We have seen a host of logics, all introduced with a claim to be constructive. Are they? What is essential for making the decision?

The second question: What exactly is the connection between the constructivist program and the phenomenon of paraconsistency? We have seen many paraconsistent logics in the course of the last chapters. Is there a specific constructive or anti-realistic motivation for paraconsistency (or even dialetheism) that is underlying all these occurrences, or is the fact that these logics do not validate Explosion merely accidental? I believe that this question is closely related to the first one.

So let us start there and ask again: What is a constructive logic?

For many, "constructive logic" and "intuitionistic logic" are simply synonymous. However, H. Wansing writes:

> In a situation where there are no clear, agreed-upon, individually necessary and jointly sufficient conditions for the constructiveness of a logical system, it seems quite difficult or next to pointless to designate one particular logic as the correct constructive logic. Nevertheless,

> for some reasons certain logics may still be regarded as constructive logics (Wansing 2008, p. 342).

What are those reasons? Is there a syntactic feature that divides constructive logics from non-constructive ones? I do not mean meta-logical properties of proof systems; there might well be definitive characteristics that divide constructive proof systems from non-constructive ones. This is a deep and important issue, but I have bracketed it all throughout the book, and I will not start to go into it now.

What I mean is this: Can we tell just from knowing which inferences are valid in a logic whether the logic is constructive or not? More specifically, is there a specific form of inference that clearly marks the divide? Which inference would that be? We have seen all Double Negation Laws and De Morgan's laws validated by N_3, surely a constructive logic. A natural assumption that is often made is that the most central and essential syntactic marker of constructivity is that failure of the Law of Excluded Middle, $\vDash A \vee \neg A$. But we have seen logics such as N_{3f} that even validate LEM. Is this a reason to deny that it is a constructive logic?

In fact, between them, the logics I discussed in Chaps. 8 and 9 validate each and every inference of classical logic. So, if they indeed deserve to be called constructive (and I think they do), then no single inference can claim to divide constructive from non-constructive logics.

However, it is hard not to have noticed that all the constructive logics that validate LEM are paraconsistent, which brings us to my second concern.

Priest takes up the question whether or not paraconsistency has a special relation to the anti-realistic/constructivist project and comes up with a negative answer: The question whether or not a logic is paraconsistent is completely independent of whether or not it is suitable to the constructivist.

> [T]here are many (...) paraconsistent logics, of widely different kinds. To determine on which side of the realism/anti-realism fence each sits requires its own investigation. Sometimes this will be obvious. For example, if the logic verifies the LEM, it is not going to sit on the anti-realist side (Priest 2012, p. 190).

As I just suggested, I think that no inference on its own, be it LEM or another one, can decide whether a logic is constructive or not. Thus a paraconsistent logic is not ruled out for constructive purposes, as Priest suggests, just because it validates LEM.

However, I am of course not claiming that the question which inferences a logic validates and fails to validate is irrelevant when we try to decide on its constructivity. The failure of LEM *for the right kind of reason* is indeed essential to the constructivity of intuitionistic logic and N_3.

What is this reason? It is the nature of the semantic values. At the very outset of this book, I have defined constructivism as the view that semantic values must be epistemically accessible. But even if such an explicit explication of the notion of constructivity in terms of semantical values is not assumed, I believe that without a semantical account, the question whether or not a logic is constructive seems quite pointless. But given the semantics, we have seen for intuitionistic logic and N_3, the failure of LEM quite distinctly signals the constructivity of the logic.

In exactly the same way, the paraconsistency of the falsificationistic logics we have seen is quite essential. And the interesting fact here is this: The reason for which LEM fails for $N3$ is exactly the same reason for which Explosion fails for $N3_f$, and this reason is the existence of *gaps* between the constructive semantic values in the semantics. After reading this book, the surprise inherent in this result may well have worn off, but it is worth taking a step back and appreciating it anew: Paraconsistency is induced, not by *gluts*, but by gaps, palpable manifestations of our epistemic humility.

If it were not for those gaps, our models would have no right to call themselves constructive models. That is, if we were to restrict our attention to models in which every statement is either verifiable or falsifiable at each world, we can no longer maintain that these senses of "verified" and "falsified" are actually constrained by our epistemic achievements and potentials any more.[1]

Now, given that the root of falsificationistic paraconsistency lies in the gaps, it is clear that *dialetheism* has no hope of gaining motivation here. If anything, what we are dealing with is a form of *analetheism*. As I presented analetheism, it is a doctrine about truth and falsity and the empty space between the two. So, for falsificationism to entail a view of this sort, it seems we need to be willing to make the identification of verifiability and truth on the one hand, and falsifiability and falsity on the other. I was not showing much enthusiasm for this identification in the last chapter. But even if we do not speak of truth and falsity, the structural similarity seems so close and clear that it would seem to me to be apt to call the falsificationistic theories cases of *constructive analetheisms*.

In sum, Priest is partly right: The question whether a logic is constructive or not is in a sense independent of the question whether it is paraconsistent or not. However, against Priest I want to contend that the sense of this independence is the same in which constructivity and the LEM are independent. Given a certain semantic setup, either LEM or paraconsistency becomes essential features of a constructive logic of one type or another.

[1] This is not meant to rule out the intuitionistic models as non-constructive. But while it is true that there are no gaps between the values 1 and 0 in these models, this is because the values are to be read as "provable/verifiable" and "not provable/verifiable," not as "verifiable" and "falsifiable."

Appendix A
Characteristics of Some of the Logics

	N_3	N_{3f}	N_{3TOL}	N_{3TOLf}	N_{AND}	N_{ANDf}	N_{OR}	N_{ORf}
$\models A \vee -A$	n	Y	n	Y	n	Y	n	Y
$-(A \vee -A) \models B$	Y	n	Y	n	Y	n	Y	n
$\models -(A \wedge -A)$	n	Y	n	Y	n	Y	n	Y
$(A \wedge -A) \models B$	Y	n	Y	n	Y	n	Y	n
$A \supset B \models -B \supset -A$	n	Y	n	Y	Y	Y	Y	Y
$-(A \supset B) \models -(-B \supset -A)$	Y	n	Y	n	Y	Y	Y	Y
$-A \models A \supset B$	Y	Y	Y	Y	n	Y	Y	Y
$B \models A \supset B$	Y	Y	Y	Y	n	Y	Y	Y
$A, A \supset B \models B$	Y	n	n	n	Y	n	n	n
$-B, A \supset B \models -A$	n	n	Y	n	Y	n	n	n
$\neg A, A \vee B \models B$	Y	n	Y	n	Y	n	Y	n
All Double Negation and de Morgan laws valid throughout								
Characteristics of IL and DIL on page 135, for HYBRID CONSEQUENCE on page 191								

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9,

Appendix B
Tableaux for N_3, N_{3f}, N_{AND}, $N_{\text{AND}f}$ and Hybrid Consequence

In this appendix, I give tableaux proof systems for the logic N_3 and the logics I introduced in the book that seem to me the most interesting and promising ones, N_{3f}, N_{AND}, $N_{\text{AND}f}$, and HYBRID CONSEQUENCE. I write N_X to talk about the two verificationistic logics, and N_{Xf} to mean the two falsificationistic ones.

The tableaux system for N_3 is taken from Priest (2008, p.176).[1] All the others are variations of that system. I prove soundness and completeness below. These systems are easy to handle and thus suitable if one's main interest is whether or not a given inference is valid, even if they might not be as interesting as other styles of proof systems for metatheoretical investigations.

Tableaux are tree structures that branch from top to bottom (unlike most other trees). The nodes of the tree are of the form $A + i$, $A - i$ or $i \leq j$, where A is a formula and i and j are natural numbers. Intuitively, these numbers represent the worlds of the Kripke models; $A + i$ means that A is verified at i. $A - i$, however, does not mean that A is falsified at i. It merely means that A is not verified at i.

$i \leq j$ means that world j is accessible from world i.

The trees for N_3 are constructed in accordance with the following rules.

[1] Note that Priest uses I_3 instead of N_3. In the book, one also finds a tableaux system for intuitionistic logic, and Priest (2009) has one for dual intuitionistic logic.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9,

Frame Rules:

$$\begin{array}{c} . \\ | \\ i \leq i \end{array} \qquad\qquad \begin{array}{c} i \leq j \\ j \leq k \\ | \\ i \leq k \end{array}$$

Heredities:

$$\begin{array}{c} p + i \\ i \leq j \\ | \\ p + i \end{array} \qquad\qquad \begin{array}{c} -p + i \\ i \leq j \\ | \\ -p + i \end{array}$$

Conjunction:

$$\begin{array}{c} A \wedge B + i \\ | \\ A + i \\ B + i \end{array} \qquad\qquad \begin{array}{c} A \wedge B - i \\ \swarrow \quad \searrow \\ A - i \quad B - i \end{array}$$

Disjunction:

$$\begin{array}{c} A \vee B + i \\ \swarrow \quad \searrow \\ A + i \quad B + i \end{array} \qquad\qquad \begin{array}{c} A \vee B - i \\ | \\ A - i \\ B - i \end{array}$$

Double Negation:

$$\begin{array}{c} - - A + i \\ | \\ A + i \end{array} \qquad\qquad \begin{array}{c} - - A - i \\ | \\ A - i \end{array}$$

de Morgan's:

$$\begin{array}{c} -(A \wedge B) + i \\ | \\ -A \vee -B + i \end{array} \qquad\qquad \begin{array}{c} -(A \wedge B) - i \\ | \\ -A \vee -B - i \end{array}$$

$$\begin{array}{c} -(A \vee B) + i \\ | \\ -A \wedge -B + i \end{array} \qquad\qquad \begin{array}{c} -(A \vee B) - i \\ | \\ -A \wedge -B - i \end{array}$$

Conditional:

$$\begin{array}{c} A \supset B + i \\ i \leq j \\ \swarrow \quad \searrow \\ A - j \quad B + j \end{array} \qquad\qquad \begin{array}{c} A \supset B - i \\ | \\ i \leq j \\ A + j \\ B - j \end{array}$$

Note: The edges represent the application of the rule. That is, in the first rule, both $A \supset B + i$ and $i \leq j$ must be present on the branch to justify the application of the rule. The rule is applied to all j s. t. $i \leq j$ occurs on the branch.

In the second rule, $i \leq j$ is added to the branch as a consequence of the application of the rule. Not only does $i \leq j$ not have to be on the branch prior to that, j must in fact be completely new to the branch. (This reflects the difference between the universal and the existential quantifier in the semantical clauses.)

$$\begin{array}{c} -(A \supset B) + i \\ | \\ A + i \\ -B + i \end{array} \qquad\qquad \begin{array}{c} -(A \supset B) - i \\ \swarrow \quad \searrow \\ A - i \quad -B - i \end{array}$$

A branch is said to be *closed* iff for some A, both $A\ + i$ and $A\ - i$ are on it, or if both $A\ + i$ and $-A\ + i$ are on it. A tree is said to be closed iff all its branches are closed.

To check the validity of an inference $A_1 \ldots A_n \models B$ in N_3, start a tree with nodes $A_1\ + 0, \ldots A_n\ + 0$, and $B\ - 0$. Apply the inference rules and see whether the tree closes. If so, the inference is valid. If not, a counter model can be read off of each open branch in the following way: For each natural number i on the branch we posit a world w_i. If $i \leq j$ is on the branch, we postulate $w_i \leq w_j$. For atomic statements, we set v such that $w_i \Vdash_1 p$ iff $p\ + i$ is on the branch, and $w_i \Vdash_0 p$ iff $-p\ + i$ is on the branch. The resulting model is said to be *induced* by the open branch.

Tableaux for the logic N_{AND} are generated in the same way, except for the rule for the (unnegated) conditional, which is:

N_{AND}:

$$\begin{array}{c} A \supset_{\mathrm{AND}} B + i \\ i \leq j \\ \swarrow \quad | \quad \searrow \\ -A + j \quad \begin{array}{c} A - j \\ -B - j \end{array} \quad B + j \end{array} \qquad\qquad \begin{array}{c} A \supset_{\mathrm{AND}} B - i \\ | \\ i \leq j \\ \swarrow \quad \searrow \\ \begin{array}{c} A + j \\ B - j \end{array} \quad \begin{array}{c} -A - j \\ -B + j \end{array} \end{array}$$

To check an inference in N_{3f} and $N_{\mathrm{AND}f}$, start the tree with nodes $-A_1\ -0, \ \ldots -A_n\ - 0$ and $-B\ + 0$, the rest stays the same.

For a tableaux system for HYBRID CONSEQUENCE, start a tree for an inference $\Gamma \mid \Delta \models_{hyb} \Phi \mid \Psi$ with nodes $A\ + 0$ for all A in Γ, $-B - 0$ for all B in Δ, $C\ - 0$ for all C in Φ and $-D\ + 0$ for all D in Ψ.

Soundness and Completeness Proofs

(Note: The metatheory in what follows is quite classical, and thus only of much use and interest for someone taking what in the body of the book I call an *eclectic* stance.)

Soundness

Definition A Nelson model $M = [W, \leq, v]$ is *faithful* to a branch b of a given tableau iff there is a mapping f from the natural numbers to W such that:

for every node $A\ +i$, $v_{f(i)}(p) = 1$
for every node $A\ -i$, $v_{f(i)}(p) \neq 1$
if $i \leq j$ is on b, then $f(i) \leq f(j)$.

Soundness Lemma *Let b be any branch of a tableau, and $M = [W, \leq, v]$ be any model. If M is faithful to b, and a tableau rule is applied to b, then it produces at least one extension, b', such that M is faithful to b'.*

This lemma is proven for N_3 in Priest (2008, p.183) . To prove it for N_{AND}, one has to go on to show that if $A \supset_{\text{AND}} B\ +i$ or $A \supset_{\text{AND}} B\ -i$ is on b and a N_{AND} model M is faithful to b, then that model is faithful to at least one of the extensions that result from applying the conditional rule to b.

The semantic clause of the conditional $\supset_{\text{AND}}$ was this:

$w \Vdash_1 A \supset_{\text{AND}} B$ iff for all $x \geq w$, $x \nVdash_0 B$ or $x \Vdash_0 A$ AND for all $x \geq w$, $x \nVdash_1 A$ or $x \Vdash_1 B$

which is equivalent to

$w \Vdash_1 A \supset_{\text{AND}} B$ iff for all $x \geq w$, $x \Vdash_0 A$ or ($x \nVdash_0 B$ and $x \nVdash_1 A$) or $x \Vdash_1 B$

Suppose that, given a mapping f, M is faithful to b and that $A \supset_{\text{AND}} B\ +i$ and $i \leq j$ occur on b. So $f(i) \Vdash_1 A \supset_{\text{AND}} B$. That means that either $f(j) \Vdash_0 A$, both $f(j) \nVdash_0 B$ and $f(j) \nVdash_1 A$, or $f(j) \Vdash_1 B$. In each of these cases, M will be faithful to one of the branches that result from an application of the first conditional rule to b.

Suppose next that, given a mapping f, M is faithful to b and that $A \supset_{\text{AND}} B\ -i$ occurs on b. This means that $f(i) \nVdash_1 A \supset_{\text{AND}} B$. From the semantic clause above and principles of first order logic, we get that there is some world v such that $f(i) \leq v$, and either both $v \nVdash_0 A$ and $v \Vdash_0 B$ or both $v \nVdash_1 B$ and $v \Vdash_1 A$. Now, take a mapping f' that is just like f, except that it assigns j to this world v. It is easy to see that under this new mapping, M will be faithful to either the left or the right branch of the expansion of b according to the rule above (as j was new to the branch, there is no danger that the new mapping f' might go wrong at worlds corresponding to earlier indices).

This establishes the soundness lemma for N_{AND}.

Soundness Theorem *for N_X (N_3, N_{AND}): For fiinte Σ, if $\Sigma \vdash A$ then $\Sigma \vDash A$.*

Proof Suppose that $\Sigma \nvDash A$. Then, there is a model M for N_X that verifies all of the premises, but not A at some world w. Let f be a function such that $f(0) = w$. This shows that M is faithful to the initial list of nodes on the tableau. From here, the soundness lemma tells us that any application of a rule will result in at least one branch to which M is faithful. If that branch were to close, then for some B, either both $B\ +i$ and $B\ -i$ would be on it, or both $B + i$ and $-B + i$ would be on it. But then we would have either$f(i) \Vdash_1 B$ and $f(i) \nVdash_1 B$, or $f(i) \Vdash_1 B$ and $f(i) \Vdash_0 B$, which is impossible in all of our models. Therefore, the branch cannot close, and $\Sigma \nvdash A$. ■

Soundness Theorem *for N_{3f} and N_{ANDf}*

Proof As the clauses and rules are not changed, the soundness lemma holds for both N_{3f} and $N_{\text{AND}f}$. The proof runs exactly parallel, except for the fact that here, if $\Sigma \nvDash A$, then all premises are unfalsified, but the conclusion is falsified. ■

Soundness Theorem for Hybrid Consequence

Again, the soundness lemma holds. The proof is as above, except that if $\Gamma \mid \Delta \nvDash_{hyb} \Phi \mid \Psi$, all premises under the BoP are verified, all others unfalsified, and the conclusion is either not verified (if it bears the BoP) or falsified (if it does not). ■

Completeness

Completeness Lemma *Let b be any open completed branch of a tableau. Let M be the model induced by b. Then, if $A + i$ is on b, then A is verified at w_i. If $A - i$ is on b, then A is not verified at w_i.*

Again, this lemma is proven for N_3 in Priest (2008, p.183) via an induction on the complexity of A. To get a proof of the lemma for N_{AND} from Priest's proof, I show that if $A \supset_{\text{AND}} B + i$ or $A \supset_{\text{AND}} B - i$ is on b and a N_{AND} model $M*$ is faithful to b, then that model verifies or, respectively, does not verify $A \supset_{\text{AND}} B$ at w_i.

If $A \supset_{\text{AND}} B + i$ is on b, then for every j such that $i \leq j$ is on b, either $-A + j$, both $A - j$ and $-B - j$, or $B + i$ is on b. Thus, by the induction hypothesis and the way the induced model is constructed, for every w_j such that $w_i \leq w_j$, either $w_j \Vdash_0 A$, both $w_j \nVdash_1 A$ and $w_j \nVdash_0 B$, or $w_j \Vdash_1 B$. Thus, $w_i \Vdash_1 A \supset_{\text{AND}} B$.

If $A \supset_{\text{AND}} B - i$ is on b , then for some j s.t. $i \leq j$, either $A + j$ and $B - j$, or $-A - j$ and $-B + j$. By construction and induction hypothesis, there is a w_j such that $w_i \leq w_j$ and either $w_j \Vdash_1 A$ and $w_j \nVdash_1 B$, or $w_j \nVdash_0 A$ and $w_j \Vdash_0 B$. Therefore, $w_i \nVdash_1 A \supset_{\text{AND}} B$.

Completeness Theorem *for N_X: For finite Σ, if $\Sigma \vDash A$ then $\Sigma \vdash A$.*

Proof Suppose $\Sigma \nvdash A$. That is, a tableau for this inference has at least one open branch. By the completeness lemmata, the model that is induced by this branch is a N_X model that verifies all premises and does not verify the conclusion. Therefore, $\Sigma \nvDash A$. ■

Completeness Theorem *for N_{Xf}: For finite Σ, if $\Sigma \vDash A$ then $\Sigma \vdash A$.*

Proof Again, as only the definition of logical consequence of these logics differs from their verificationistic counterparts, the completeness lemmata carry over without need for modification. The proof of the completeness theorem is also quite similar, only that the induced model falsifies none of the premises, but does falsify the conclusion, which means that $\Sigma \nvDash A$, as required. ■

Completeness Theorem *for Hybrid Consequence: Here, the induced model verifies all premises that bear the BoP, falsifies none of those that do not, and either does not verify the conclusion (if it is under the BoP) or falsifies the conclusion (if it is not). Thus,* $\Gamma \mid \Delta \nvDash_{hyb} \Phi \mid \Psi$. ■

References

Anderson, A., Belnap, N., Dunn, J., et al. (1992). *Entailment: The logic of relevance and necessity* (Vol. II). Princeton: Princeton University Press.

Auxier, R., & Hahn, L. (2007). *The philosophy of Michael Dummett*. Chicago: Open Court Publishing Company.

Bailenson, J., & Rips, L. (1996). Informal reasoning and burden of proof. *Applied Cognitive Psychology*, *10*(7), 3–16.

Bailenson, J. (2001). Contrast ratio: Shifting burden of proof in informal arguments. *Discourse Processes*, *32*(1), 29–41.

Beall, J. (2006). Intuitionism, realism, relativism and rhubarb. *Truth and realism* (pp. 61–74). USA: Oxford University Press.

Beall, J., & Ripley, D. (2004). Analetheism and dialetheism. *Analysis*, *64*(1), 30–41.

Beall, J., & Restall, G. (2006). *Logical pluralism*. Oxford: Oxford University Press.

Belnap, N. (1977). A useful four-valued logic. In J. Dunn & G. Epstein (Eds.), *Modern uses of multiple-valued logic* (pp. 8–37). Boston: Reidel Publishing Company.

Belnap, N. (1977). How a computer should think. In G. Ryle (Ed.), *Contemporary aspects of philosophy* (pp. 30–56). Stockfield: Oriel Press.

van Benthem, J. (2008). Logic and reasoning: Do the facts matter? *Studia Logica*, *88*(1), 67–84.

Berto, F. (2007). *How to sell a contradiction: The logic and metaphysics of inconsistency*. London: College Publications.

Blamey, S. (1986). Partial logic. In D. Gabbay (Ed.), *Handbook of philosophical logic* (Vol. 3, pp. 1–70). Dordrecht: Reidel.

Braver, L. (2007). *A thing of this world: A history of continental anti-realism*. Evanston: Northwestern University Press.

Burgess, J. (2009). *Philosophical logic*. Princeton: Princeton University Press.

Buss, S. (Ed.). (1998). *Handbook of proof theory*. Amsterdam: Elsevier.

Carnap, R. (1959). *The logical syntax of language*. Boston: Littlefield, Adams and Co.

Cogburn, J. (2004). The philosophical basis of what? The anti-realist route to dialetheism. In G. Priest, J. C. Beall, & B. P. Armour-Garb (Eds.), *The law of non-contradiction* (Vol. 1, pp. 217–235). Oxford: Oxford University Press. (Chapter 1).

Cornish, K. (1998). *The jew of linz: Wittgenstein*. Century: Hitler and their Secret Battle for the Mind.

van Dalen, D. (2004). Kolmogorov and Brouwer on constructive implication and the ex falso rule. *Russian Math Surveys*, *59*, 247–257.

Dummett, M. (1973). *Frege: Philosophy of language*. London: Duckworth.

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9,

Dummett, M. (1975). The philosophical basis of intuitionistic logic. *Studies in Logic and the Foundations of Mathematics*, *80*, 5–40.

Dummett, M. (1978). *Truth and other enigmas*. Harvard: Harvard University Press.

Dummett, M. (1987). Reply to Prawitz. In B. Taylor (Ed.), *Michael Dummett: Contributions to philosophy* (pp. 117–65). Dordrecht: Nijhoff Publishers.

Dummett, M. (1991). *The logical basis of metaphysics*. Harvard: Harvard University Press.

Dummett, M. (1993). *The seas of language*. Oxford: Clarendon Press.

Dummett, M. (1998). Truth from the constructive standpoint. *Theoria*, *64*(2–3), 122–138.

Dummett, M. (2000). *Elements of intuitionism*. USA: Oxford University Press.

Dummett, M. (2004). *Truth and the past*. New York: Columbia University Press.

Dummett, M. (2007). Reply to Ian Rumfitt. In R. Auxier (Ed.), *The philosophy of Michael Dummett* (pp. 694–698). Chicago: Open Court Publishing Company.

Fitting, M. (2002). Bilattices are nice things. In T. Bolander, et al. (Eds.), *Self-reference*. Stanford: CSLI Publications.

Garcia-Carpintero, M., & Kölbel, M. (Eds.). (2008). *Relative truth*. Oxford: Oxford University Press.

Goodman, N. (1981). The logic of contradiction. *Zeitschrift fuer mathematische Logik und Grundlagen der Mathematik*, *27*, 119–126.

Green, K. (2001). Dummett: Philosophy of language. Polity: Wiley.

Green, K. (2005). The context principle and Dummett's argument for anti-realism. *Theoria*, *71*(2), 92–117.

Hale, B. (1999). Realism and its oppositions. In B. Hale & C. Wright (Eds.), *A companion to the philosophy of language*. New York: Wiley.

Halmos, P., & Givant, S. (1998). *Logic as algebra*. Number 21. Mathematical Association of America.

Horn, L. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Jaśkowski, S. (1969). Propositional calculus for contradictory deductive systems. *Studia Logica*, *24*(1), 143–157.

Jaspars, J. (1994). *Calculi for Constructive Communication: A Study of the Dynamics of Partial States*. Language and Computation: Institute for Logic.

Jaspars, J. (1995). Partial up and down logic. *Notre Dame Journal of Formal Logic*, *36*(1), 134–157.

Johansson, I. (1936). Der minimalkalkul, ein reduzierter intuitionistischer formalismus. *Compositio Mathematica*, *4*, 119–136.

Kapsner, A. (2013). From semantics to logic: The scenic route. Teorema: Revista internacional de filosofía, 32(1), 77–94.

Kirkham, R. (1995). *Theories of truth: A critical introduction*. Cambridge: The MIT Press.

Kölbel, M. (2009). The evidence for relativism. *Synthese*, *166*(2), 375–395.

von Kutschera, F. (1969). Ein verallgemeinerter Widerlegungsbegriff für Gentzenkalküle. *Archive for Mathematical Logic*, *12*(3), 104–118.

Laudan, L. (2006). *Truth, error and criminal law: An essay in legal epistemology*. Cambridge: Cambridge University Press.

Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, *8*(1), 339–359.

López-Escobar, E. (1972). Constructions and negationless logic. *Studia Logica*, *30*(1), 7–19.

Marcos J (2011) The value of the two values. In: Béziau J-Y, Coniglio ME (Eds) Logic without Frontiers: Festschrift for Walter Alexandre Carnielli on the occasion of his 60th birthday, vol 17. Tribute series. London, College Publications, pp 277–294

Martino, E., & Usberti, G. (1994). Temporal and atemporal truth in intuitionistic mathematics. *Topoi*, *13*(2), 83–92.

McLaughlin, R. (1990). *On the logic of ordinary conditionals. SUNY series in logic and language*. New York: State University of New York Press.

Miller, D. (2006). Making falsification bite. In Out of error: Further essays on critical rationalism (Chapter 13, pp. 249–262). Farnham: Ashgate Publishing.

Miller, A. (2002). What is the manifestation argument? *Pacific Philosophical Quarterly*, *83*, 352–383.
Miller, A. (2003). What is the acquisition argument? In A. Barber (Ed.), *Epistemology of language* (pp. 459–495). Oxford: Oxford University Press.
Monk, R. (1998). Review of "The jew of linz". Quadrant. *42*(9), 79–82.
Mortensen, C. (1995). *Inconsistent mathematics*. Berlin: Springer.
Nelson, D. (1949). *Constructible falsity. Journal of Symbolic logic*, *14*(1), 16–26.
Odintsov, S. (2008). *Constructive negations and paraconsistency*. Berlin: Springer.
Pietz, A. (2011a). The logic of falsification. In M. Peliš, V. Punčochář (Eds.), *The logica yearbook 2010* (pp. 169–180). London: College Publications.
Pietz, A. (2011b). Not quite intuitionism. In XXII. Deutscher Kongress für Philosophie. http://epub.ub.uni-muenchen.de/12589/.
Pietz, A., & Rivieccio, U. (2013). Nothing but the truth. *Journal of Philosophical Logic*, *42*(1), 125–135.
Placek, T. (1999). *Mathematical intuitionism and intersubjectivity: A critical exposition of arguments for intuitionism*. Dordrecht: Kluwer Academic Publishers.
Popper, K. (1948). On the theory of deduction, parts I & II. Indagationes Mathematicae, 10(1–2), 44–54, 111–120.
Prawitz, D. (1987). Dummett on a theory of meaning and its impact on logic. In B. Taylor (Ed.), *Michael Dummett: Contributions to philosophy* (pp. 117–65). Dordrecht: Nijhoff Publishers.
Priest, G. (2009). *Dualising intuitionistic negation. Principia*, *13*, 165–189.
Priest, G. (2003). Paraconsistent logic. In D. Gabbaz & F. Guenther (Eds.), *Handbook of philosophical logic*. Dordrecht: Kluwer Academic Publishers.
Priest, G., Beall, J. C., & Armour-Garb, B. P. (Eds.). (2004). *The law of non-contradiction: New philosophical essays*. USA: Oxford University Press.
Priest, G. (2006). *In contradiction: A study of the transconsistent*. USA: Oxford University Press.
Priest, G. (2006). *Doubt truth to be a liar*. USA: Oxford University Press.
Priest, G. (2008). *An introduction to non-classical logic: From if to is*. Cambridge: Cambridge University Press.
Priest, G. (2012). Realism, antirealism, and paraconsistency. In M. Rahman & G. Primiero (Eds.), *The realism-antirealism debate in the age of alternative logics* (pp. 181–190). Berlin: Springer.
Rasiowa, H. (1974). *An algebraic approach to non-classical logics*. Amsterdam: North-Holland Publishing Company.
Rauszer, C. (1980). An algebraic and Kripke-style approach to a certain extension of intuitionistic logic. (Vol. CLXVII of Dissertationes Mathimaticae. Institute of Mathematics, Polish Academy of Sciences 1980).
Rauszer, C. (1977). Applications of Kripke models to Heyting-Brouwer logic. *Studia Logica*, *36*(1), 61–71.
Restall G. (1997). Extending intuitionistic logic with subtraction. www.consequently.org/writing/extendingj.
Rosenkranz, S. (2002). *Objectivity and Realism*. Synchron: Meeting the Manifestation Challenge.
Rumfitt, I. (2007). Asserting and excluding: Steps towards an anti-realist account of classical logic. In R. Auxier (Ed.), *The philosophy of Michael Dummett* (pp. 639–693). Chicago: Open Court Publishing Company.
Schlick, M. (1936). Meaning and verification. *The Philosophical Review*, *45*, 339–369.
Schroeder-Heister, P. (2009). Schluss und Umkehrschluss. In Lebenswelt und Wissenschaft; Deutsches Jahrbuch Philosophie, Bd.3. Bouvier.
Shapiro, S. (2000). *Thinking about mathematics*. Oxford: Oxford University Press.
Shieh, S. (1998). Undecidability in anti-realism. *Philosophia Mathematica*, *6*(3), 324–333.
Shramko, Y. (2005). Dual intuitionistic logic and a variety of negations: The logic of scientific research. *Studia Logica*, *80*(2), 347–367.
Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory*. Oxford: Blckwell.

Stalnaker, R. (1984). *Inquiry*. Cambridge: Bradford Books.

Stojanovic, I. (2007). Talking about taste: Disagreement, implicit arguments, and relative truth. *Linguistics and Philosophy*, *30*(6), 691–706.

Strawson, P. (1950). On referring. *Mind*, *59*(235), 320–344.

Tennant, N. (2002). *The taming of the true*. Oxford: Oxford University Press.

Troelstra, A. (1990). On the early history of intuitionistic logic. *Mathematical logic* (pp. 3–17). Berlin: Plenum Press.

Urbas, I. (1996). Dual-intuitionistic logic. *Notre Dame Journal of Formal Logic*, *37*(3), 440–451.

Vakarelov, D. (2006). Non-classical negation in the works of helena rasiowa and their impact on the theory of negation. *Studia Logica*, *84*(1), 105–127.

Wansing, H. (1993). The logic of information structures. In Lecture notes in AI 681. Berlin: Springer.

Wansing, H. (2010). Connexive logic. Stanford Encyclopedia of Philosophy, Fall: http://plato.stanford.edu/archives/fall2010/entries/logic-connexive/.

Wansing, H. (1998). *Displaying modal logic*. Dordrecht: Kluwer Academic Publishers.

Wansing, H. (2008). Constructive negation, implication, and co-implication. *Journal of Applied Non-Classical Logics*, *18*, 341–364.

Wansing, H. (2012). A non-inferentialist, anti-realistic conception of logical truth and falsity. *Topoi*, *31*(1), 93–100.

Williamson, T. (1994). Never say never. *Topoi*, *13*(2), 135–145.

Wright, C. (2006). Intuitionism, realism, relativism and rhubarb. *Truth and realism* (pp. 38–60). USA: Oxford University Press.

Wright, C. (1993). *Realism, meaning and truth*. Oxford: Blackwell.

Index

A
Accommodation, 187
Acquisition argument, 22
Analetheism, 76, 115, 201
Anti-realism, 13
Assertibility, 22
Assertoric content, 69, 96
Ayer, A.J., 90

B
Beall, J.C., 27, 76, 115
Belnap, N., 59, 121
Bivalence, 14, 17, 21, 68, 75, 97
Brouwer, L., 31, 128
Brouwer–Heyting–Kolmogorov interpretation, 34
Burden of proof, 112
  strategy, 102, 185

C
Carnap, R., 27
Co-implication, 41, 128, 132, 141, 142
Cogburn, J., 151
Commands, 106
Conditional logic, 157
Conditionals, 120, 137, 157
Constructible falsity property, 154
Constructive logic, 13, 19, 26, 94, 199
Constructivism, 12
Contraposable consequence, 156
Contraposition, 154, 155, 157
Conversational score, 187
Correct assertibility, 110
Correctness, 118
CV&IF strategy, 97, 176

D
Da Costa, N., 128
Decidability, 15
Designated values, 61, 69, 192
Dialetheism, 57, 72, 167, 201
Discourse separation strategy, 101, 175
Disjunctive property, 154
Disquotational schema, 119
Double Negation Elimination, 18, 37, 93, 135, 154
Double Negation Introduction, 18, 135, 154
Dual intuitionistic logic, 125
Dual minimal logic, 144
Duality, 75
Dummett
  on assertibility gaps, 97, 177
  on designated values, 69
  on falsificationism, 100
  on hybrid strategies, 96, 177

E
Eclecticism, 26
Empirical absurdity, 90, 91
Empty promise conversion, 39, 96
Exactly true logic, 64, 152
Expanded falsificationism, 87, 163
Expanded verificationism, 87, 145
Explosion, 38, 57, 65, 71

F
Falsifiability, 89
Falsification, 109
Falsificationism, 12, 86, 99
Faultless disagreement, 113
First degree entailment, 59

A. Kapsner, *Logics and Falsifications*,
Trends in Logic 40, DOI: 10.1007/978-3-319-05206-9,

G
Gaps, 57, 64, 67, 177, 201
Gentzen, G., 33
Givant, S., 76
Gluts, 57, 64, 150, 201

H
Halmos, P., 76
Heyting, A., 32, 38, 129
Heyting–Brouwer logic, 129, 142
Hybrid consequence, 190
Hybrid strategies, 87, 96, 175
Hybrid strategy, 101

I
Incoherence, 167
Incorrectness, 106
Ingredient sense, 69, 96
Intuitionism, 13

J
Jaśkowski, S., 170

K
Kleene, S., 67
Kolmogorov, A., 38

L
Laudan, L., 185
Law of Excluded Middle, 17, 97, 135, 165, 200
Law of Excluded Third, 18, 33, 37, 97, 154
Law of non-contradiction, 72, 75, 135
Legal discourse, 112
Lewis, D., 157, 187
Logical pluralism, 27
Logical positivism, 13
Lopez-Escobar, E., 147
Łukasiewicz, J., 67

M
Manifestation argument, 21
Mannoury, G., 32, 129
Meaning, 19
  as use, 20
Meaning theory, 24
Metaphysics, 13, 19
Miller, D., 128
Minimal logic, 39, 46, 71, 92
Modal coherence
  elaborate, 171
  simple, 171
Modus ponens, 166
Modus tollens, 140, 154, 166
Monotonicity, 130

N
N_3, 147
N_4, 147
Negation
  dual intuitionistic, 143
  intuitionistic, 35, 159
    defined in N_3, 153
  toggle, 93, 159
Negation incompatibility principle, 72, 136
Nelson, D., 147
No gluts principle, 72, 136
Non-monotonic logic, 152

P
Paraconsistency, 71, 135, 165, 200
Philosophy of language, 19
Pietz, A., 64
Popper, K., 89, 129
Presumption of innocence, 113, 185
Priest, G., 66, 74, 171
Provability, 47, 88
Pure falsificationism, 87, 125
Pure verificationism, 87

R
Relevant logic, 58
Restall, G., 27
Retraction, 108, 178
Revisionism, 26
Ripley, D., 76
Rivieccio, U., 64
Rumfitt, I., 127, 130, 172
Russell, B., 67

S
Safe assertibility, 127, 172
Schlick, M., 89
Schroeder-Heister, P., 133
Semantic theory, 24
Skeletal semantics, 36
Stalnaker, R., 157
Strawson, P., 68, 70

T
Tertium non datur, 18, 69, 97
Theory of meaning, 24
Truly constructive conversion, 39, 95
Truth, 22, 118, 192
  conditions, 22

U
Undeciability, 17
Unfalsifiability, 110, 115
  pro tempore, 89, 111, 121, 127
  recognizably eternal, 89, 111, 127

V
Verifiability, 88, 89
Verificationism, 12, 85
Vienna circle, 90

W
Wansing, H., 147, 170
Wright, C., 113