MICHAEL DE

Negation in Context

Thesis submitted for the degree of PhD in Philosophy
University of St Andrews, Scotland, May 2011

Only few wind up where they expected
The rest never know where they are headed

Abstract

The present essay includes six thematically connected papers on negation in the areas of the philosophy of logic, philosophical logic and metaphysics. Each of the chapters besides **Chapter 1**, which puts each the chapters to follow into context, highlights a central problem negation poses to a certain area of philosophy. **Chapter 2** discusses the problem of logical revisionism and whether there is any room for genuine disagreement, and hence shared meaning, between the classicist and deviant's respective uses of 'not'. If there is not, revision is impossible. I argue that revision is indeed possible and provide an account of negation as contradictoriness according to which a number of alleged negations are declared genuine. Among them are the negations of **FDE** (First-Degree Entailment) and a wide family of other relevant logics, **LP** (Priest's dialetheic "Logic of Paradox"), Kleene weak and strong 3-valued logics with either "exclusion" or "choice" negation, and intuitionistic logic. **Chapter 3** discusses the problem of furnishing intuitionistic logic with an empirical negation for adequately expressing claims of the form 'A is undecided at present' or 'A may never be decided' the latter of which has been argued to be intuitionistically inconsistent. **Chapter 4** highlights the importance of various notions of consequence-as-s-preservation where s may be falsity (versus untruth), indeterminacy or some other semantic (or "algebraic") value, in formulating rationality constraints on speech acts and propositional attitudes such as rejection, denial and dubitability. **Chapter 5** provides an account of the nature of truth values regarded as objects. It is argued that only truth exists as the maximal truthmaker. The consequences this has for semantics representationally construed are considered and it is argued that every logic, from classical to non-classical, is gappy. Moreover, a truthmaker theory is developed whereby only *positive* truths, an account of which is also developed therein, have truthmakers. **Chapter 6** investigates the definability of negation as "absolute" impossibility, i.e. where the notion of necessity or possibility in question corresponds to the global modality. The modality is not readily definable in the usual Kripkean languages and so neither is impossibility taken in the broadest sense. The languages considered here include one with counterfactual operators and propositional quantification and another bimodal language with a modality and its complementary. Among the definability results we give some preservation and translation results as well.

Acknowledgements

I remember vibrantly my first years as an undergraduate in philosophy. Since philosophy was not offered as part of my high school curriculum, this was my first taste of the *real* thing. How sweet it was. During that time I was influenced by two people who were responsible for steering me in the direction I lie today. Those two incisive philosophers are David DeVidi and Tim Kenyon. David enthusiastically introduced me to the wonderful world of logic. Tim acted as my Honors thesis advisor and we met weekly, early, and even sometimes on weekends. I recall the innumerable hours I spent in the library preparing for those meetings. At the time I thought I was doing *serious* research. How I've learned!

There are two other incisive philosophers to whom I would like to express my deepest gratitude. The first of those is Jeff Pelletier, my Masters advisor. His continued support and guidance through the beginnings of my research career had a significant impact on me as both a philosopher and a person. The second is Stephen Read, my doctoral thesis advisor, who served to both inspire and guide. Our weekly meetings in the Arché Logic Group, The Foundations of Logical Consequence and Medieval Logic Group seminars were highly stimulating and enjoyable. Much of his influence can be seen in nearly every page of this essay.

During my time at the Arché Philosophical Research Centre, I had the pleasure of countless enjoyable exchanges with some very bright minds. Among those, I would like to thank Ralf Bader, Derek Ball, Björn Brodowski, Herman Cappelen, Colin Caret, Yuri Cath, Paul Dimmock, Dylan Dodd, Roy Dyckhoff, Philip Ebert, Noah Friedman-Biglin, Patrick Greenough, Katherine Hawley, Andri Hjálmarsson, Thomas Hodgson, Torfinn Huvenes, Ole Hjortland, Jonathan Ichikawa, Frederique Janssen-Lauret, Dirk Kindermann, Ira Kiourti, Toby Meadows, Julien Murzi, Dilip Ninan, Andrea Onofri, Walter Pedriali, Laura Porro, Graham Priest, David Ripley, Marcus Rossberg, Gil Sagi, Anders Schoubye, Daniele Sgaravatti, Stewart Shapiro, Martin Smith, Andreas Stokke, Margot Strohminger, Crispin Wright, and Elia Zardini.

I would like to express special gratitude to Julia Langkau who, among endless other generosities, proof read a draft of the thesis and quarreled with me uncompromisingly over points of style and presentation.

Thanks is owed to the Institute for Logic, Language and Computation at the University of Amsterdam for their warm hospitality during the winter months of 2009 and, in particular, to Johan van Benthem, Catarina Dutilh Novaes, Dick de Jongh, Alexandru Morcoi, Sarah Uckelman, Jouko Väännen, and Yde Venema for making the visit most highly enjoyable. Thanks also to the St Andrews/Stirling Philosophy Graduate Program and the Philosophical Quarterly

who, in the first place, made this visit possible.

Many others have had a positive impact on me in other ways, and a final thanks is owed to them as well.

Contents

Chapter 1

Negation in Context

Abstract

This chapter sets up the background for the chapters to follow. It includes historical references, puts the problems discussed in the forthcoming chapters into context, and ties together a number of the main themes and problems.

1.1 Introduction

Plato's beard and contemporary issues

When one takes a close look at "new" problems in philosophy one may often find that they are really old problems in modern guise. This is not a bad thing. The old problems that are truly central to philosophy take on new form and become amenable to solutions involving a host of new tools and insights. There may be a lot to learn by viewing the same problem in different contexts. It is important, however, that one sees the connection between the old and the new so that an appreciation of the insights garnered centuries, indeed thousands, of years ago may likewise be appreciated in modern light.

Some of the problems that I have in mind concern negation and falsity. For instance, Plato's discussion (in the *Sophist*, *Theaetetus*, and *Parmenides*) of the problem of speaking about *that which is not*, a problem labeled *Plato's beard* by Quine in [W. 61][1], recurred in numerous forms throughout history. The problem was arguably the primary motivation of Russell's theory of descriptions, a theory which provided a more ontologically amenable solution to

---

[1]It is claimed, e.g. by Read [Rea94, ch. 5], that Quine actually used the label 'Plato's beard' to denote what he thought was Plato's solution to what I've called the problem. But it's not clear whether Quine meant the problem or a particular solution. Quine says "This is the old Platonic riddle of nonbeing. Nonbeing must in some sense be, otherwise what is it that there is not? This tangled doctrine might be nicknamed Platos beard; historically it has proved tough, frequently dulling the edge of Occam's razor" [W. 61, pp. 1–2]. Notice that he speaks of the "doctrine" as a "riddle" and says that it has proved "tough" which better applies to problems than solutions. In any case, the point is merely terminological. Like Read, I prefer to call the problem 'Plato's beard' even if that is not how Quine used it.

Plato's beard than Russell's earlier solution which distinguished existence from being, attributing only the latter to Pegasus. Today the problem takes shape as that of finding truthmakers for negative truths, in particular, for negative existentials. There is no doubt it will take on yet other shapes in the future.

Let us take a look at two particular cases, one already alluded to, where the old and new coincide; the old being Plato's beard and the new coming from two disparate areas of philosophy—contemporary metaphysics, on the one hand, and constructive mathematics on the other. The two new turn out to be the one old. But first let me briefly describe the old, Plato's beard, before the connections are drawn.

Plato's beard concerns the puzzle of how we can make sense of statements that involve so-called empty names. For if the application of a predicate to a term is to make any sense then, as the reasoning goes, it must be presupposed that the term refers, i.e. to an existing entity. For example, how can 'Pegasus does not exist' have sense if 'Pegasus' does not refer? Are we not then speaking of nothing, and isn't that equivalent to not speaking at all, or at best, saying something without meaning? Plato struggled with this problem in various places and it has been attributed to him (contra [Dur98]) that because of it he was led to the view that Pegasus has *being* in some sense. Prior to his theory of descriptions, Russell was led to the same conclusion on precisely the same grounds.

Now consider the problem negation posed for the constructive mathematician G. F. C. Griss.[2] The problem, according to Griss, is that from a constructivist point of view, it makes *no sense* to speak of that which is false and hence it makes no sense to deny truths, at least if we assume that denying truths is equivalent to asserting falsehoods. For Griss, a statement has meaning only insofar as *each* of its terms refers. Notice that this problem regarding the reference of terms is a constructivist rehashing of Plato's beard. However for Griss, Quine and Russell's purported solution to the problem will not work. Their solution was to analyze singular terms as definite descriptions and to thereby eliminate empty singular terms such as 'Pegasus' in favor of empty predicates such as 'Pegasizes', e.g. the sentence 'Pegasus does not exist' being parsed as 'Nothing Pegasizes'. This solution does not work for Griss since he regards empty predicates as meaningless on the grounds that the null class, their purported extension, fails to exist.

Griss' views led to negationless mathematics, a reconstruction of a significant part of constructive mathematics without negation. The motivation for the strong negation of Nelson (which he called 'constructible falsity') may be seen as deriving from Griss' stringent requirement that truth depend on a *positive* construction and hence not on any construction involving an absurdity, a notion in terms of which intuitionistic negation is typically defined.

Finally, a more obvious connection between old and new is found in contemporary metaphysics concerning the problem of finding truthmakers for negative truths. Russell grappled with a close cousin of this problem, eventually— though cautiously—caving in and accepting the existence of "negative facts". He amusingly stated:

> When I was lecturing on this subject at Harvard [in 1914] I argued
> that there were negative facts, and it nearly produced a riot: the

---

[2]See [Gri46].

> class would not hear of there being negative facts at all. [Rus10, p. 42]

But he was cautious. He goes on:

> I am still inclined to think that there are [negative facts]. However, one of the men to whom I was lecturing at Harvard, Mr. Demos, subsequently wrote an article in Mind to explain why there are no negative facts...I think he makes as good a case as can be made for the view that there are no negative facts. It is a difficult question. I really only ask that you should not dogmatize. I do not say positively that there are, but there may be. [Rus10, p. 42]

The problem of finding truthmakers for negative truths, i.e. truths equivalent to negations, is roughly this. If truth is grounded in the world, that is, for every truth there corresponds a truthmaker, then even negative truths have truthmakers. But exactly what could count as a truthmaker for truths such as 'Pegasus does not exist'? Surely not Pegasus, since he does not exist. Nor does any other wordly entity seem a good candidate *unless* among those entities are counted things such as negative facts (Russell), reified absences (e.g. [Kuk06]), totalities (Armstrong), or any of the other myriad ontological repugnances that have been posited for solutions to the truthmaker worry. These types of solution give being to nonbeing. They are solutions to yet another recasting of Plato's beard (especially as the problem appears in the *Sophist*) that Quine so vehemently rejected.

Outline

These and other problems concerning negation are the topic of this essay. Each of the subsequent chapters brings to light some significant problems negation poses to different areas of contemporary philosophy, primarily to current debates in philosophical logic, the philosophy of logic and metaphysics.

The rest of this chapter is sectioned into headings corresponding to the chapters to follow:

2. What is negation?

3. Constructivism and empirical negation

4. Rejection, denial and other negative speech acts and attitudes

5. Negation, truth and falsity

6. The definability of negation as impossibility

The following sections of the present chapter are intended to give the reader some introductory background to the chapter to which the section corresponds. When possible, I have attempted to avoid overlap between this chapter and the others.

1.2 What is negation?

Negation enjoys a rich history. Various proposals of the notion, as it occurs in various forms from subsentential to sentential, have been put forward which raises an important question: Do any of all of these proposals, or the proposed "negations", have some core features in common? If they do not, then their similarities, if they have any, are at best family resemblances. That may not seem so bad. The problem, however, is that forming a mere family is not enough to overcome the Quinean "change of subject" objection. According to Quine, as soon as the deviant's 'not' is non-classical, she is no longer expressing negation, and if she is no longer expressing negation then the deviant and classicist are merely speaking past each other when they purport to disagree about negation; for one is talking about negation and the other about "negation". The Quinean objection is simple and yet strikes at the core of logical revisionism, the idea that a theory of logical consequence is revisable.

I regard as revisable a theory of correct inference and not the object of that theory. I call the former 'logic' and the latter 'LOGIC'. The two are important to distinguish. For instance, early Quine held that "no statement is immune to revision" ([Qui51, p. 40]) and, in particular, that "[r]evision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics" ([Qui51, p. 40]). While later Quine held, apparently to the contrary, that when the deviant logician "tries to deny the doctrine he only changes the subject" [Qui70, p. 81]. One way to interpret earlier Quine consistently with later Quine is to take 'logic' in the mouth of the latter as meaning LOGIC (hence unrevisable) and in the mouth of the former as meaning logic (hence revisable).

It is then absolutely crucial to the project of logical revisionism, if it is to even have a starting chance, that disagreement between classicist and deviant be possible. One way to accomplish this is to argue that there is enough shared meaning or content in their respective uses of the relevant vocabulary to allow for disagreement. In particular, if the only relevant vocabulary is negation, then all that has to be argued for is that there is enough shared meaning in their respective uses of 'not'. In fact, in a number of important cases one may view the difference between classical and deviant, such as the intuitionist, as one concerning solely negation.[3] But even if there is more to their differences than negation, one must still account for shared content between their respective uses of 'not'.

It is commonly held that a contradictory-forming operation is a negation. There is no reason to think that contradictory-forming operations are unique. Indeed, on a compelling view of contradictoriness, defined as the intersection of contrariety and subcontrariety, the classical and deviant assign the same meaning to 'not'. One may even hold the stronger view that each logic gives rise to different yet *genuine* contradictory-forming operations and hence to different negations. This sort of view might be held by a logical pluralist. One account of contradictoriness according to which the meaning of 'not' is shared by deviant and classicist is defended in chapter 2 and compared with other popular accounts that fare poorly from a deviant's perspective.

---

[3] A suggestion to this effect may be found in section 1.3.

1.3 Constructivism and empirical negation

Introduction

The difference between classical and intuitionistic *propositional* logic may be summed up as follows. Let each logical connective (e.g. $\wedge$, $\rightarrow$, etc.) be associated with what might be considered its most *natural* family of rules consisting of its introduction and elimination rules.[4] Suppose the most natural family of rules associated with a given connective are just those of Gentzen's system **NK**. Then the *only* difference between classical and intuitionistic logic is that only the former is associated with the rule of double negation elimination and only the latter with *ex falso quodlibet* (EFQ).[5] In other words, the difference between the two logics is merely a difference of negation.

This is not the only way to draw the difference between the two rivals. For instance, Read [Rea00, pp. 144–148] argues that the difference between the two logics is not one of negation but rather one of implication, with (the rules for) classical implication being stronger than those of its intuitionistic rival. Indeed there are numerous ways to draw the difference between the two (e.g. one might also point to disjunction), but drawing it in terms of negation seems more natural than drawing it in terms of any other connective, at least from the point of view of natural deduction and also, in a number of cases, from a semantical point of view as well.

Read [Rea00] disagrees, arguing that the difference between intuitionistic and classical logic has to do with the conditional. He holds that a correct natural deduction system formulation of classical logic is one which strengthens rules for the conditional by allowing "disjoined parameters", essentially a multiset of formulae to be thought of as a disjunction of its elements, to occur in the conclusion of the introduction rule $\rightarrow I$ of the conditional.[6] He is therefore endorsing multiple conclusions for natural deduction either implicitly or explicitly. However there is much to be said for the naturalness of single conclusions and the unnaturalness of multiple conclusions (on which see, e.g., [Rum00, pp. 795–796]), whether multiple conclusions are disguised or not.[7] Multiple conclusions are also objected to for reasons not associated with naturalness; see, e.g. [Ten97], which is criticized implicitly by [Rea00, p. 145] as begging the question, who claims that multiple conclusions smuggle in the law of excluded middle through the back door. This may be taken as reason enough for not preferring to draw the distinction between intuitionistic and classical logic in terms of the conditional, at least insofar as the distinction is

---

[4] I say 'family' instead of 'pair' since rules for a connective typically do *not* come in pairs, e.g. $\vee$ is often associated with a pair of introduction rules and a single elimination rule. Of course, this depends on what one means by 'rule'. For instance, we might think of $A_1 \wedge A_2 \vdash A_i$ for $i \leq 2$ as a *single* elimination rule for $\wedge$ but such an interpretation of 'rule' would be nonstandard though not entirely absent from the literature (e.g. see fn 18 of chapter 6).

[5] Of course if we consider derived rules along with the primitive ones then classical negation too will be associated with EFQ.

[6] As a matter of convention I shall use $\otimes I$ and $\otimes E$ to denote the introduction and elimination rules for a connective in the setting of natural deduction. In the usual cases, e.g. as with $\wedge$ and $\vee$, when there is more than one introduction and elimination rule for a given connective, I shall use $\otimes I$ or $\otimes E$ to ambiguously denote either one when the ambiguity is harmless. For instance $\wedge E$ may denote either $A \wedge B \vdash A$ or $A \wedge B \vdash B$.

[7] They may be thought of as disguised in [Kne56].

drawn in [Rea00]. There may be less contentious ways of drawing it in terms of
the conditional that, for example, do not appeal to multiple conclusion natural
deduction systems.

From a semantic perspective, we also see the distinction being drawn in
terms of negation. Algebraic semantics for intuitionistic logic (in particular
when the class of algebras under consideration is the class of Heyting algebras)
and classical logic marks the distinction between the two in terms of comple-
mentation, the operation assigned to their respective negations. In particular,
Heyting algebras are a generalization of boolean algebras in the sense that
boolean complementation must satisfy more properties than the "pseudocom-
plementation" of Heyting algebras.

Precisely, a Heyting algebra is a distributive lattice $(A, \wedge, \vee, \bot, \top)$ (where
$\bot$ and $\top$ are the bottom and top elements, respectively) satisfying the property
that for all $a, b \in A$ there exists an element $a \rightarrow b \in A$ such that for every
$c \in A$ we have

$$c \leq a \rightarrow b \iff a \wedge c \leq b.$$

We can define the *pseudocomplement* $\neg a$ of a as $a \rightarrow \bot$. We call the just
defined complementation operation, $\neg$, *pseudocomplementation*.

A boolean algebra is a Heyting algebra $(A, \wedge, \vee, \bot, \top)$ whose complementa-
tion $\neg$, called *boolean complementation*, also defined as $a \rightarrow \bot$ for $\neg a$, satisfies
the following:

$$a \wedge \neg a = \bot \qquad\qquad a \vee \neg a = \top.$$

Clearly every complementation is a pseudocomplementation but not conversely,
and so the class of boolean algebras is a subclass of the class of Heyting algebras.

One sees the analogous result holding for Kripke semantics for classical
and intuitionistic logic where the class of "sheer reflexive frames", i.e. frames
where the accessibility relation is the identity relation on the class of worlds,
characterizing classical logic is a subclass of the class of preorders character-
izing intuitionistic logic.[8] However the difference between Kripke semantics
for classical and intuitionistic logic does not obviously have anything to do
with negation since the truth conditions for negation on both semantics are
the same, viz. those given by

$$M, a \models \neg A \iff \forall b \geq a, M, b \not\models A.$$

The only difference is that the semantics is given relative to different classes
of structures, e.g. the class of preorders for intuitionistic logic and the class of
sheer reflexive frames for classical logic.

In any case, whether the difference between the two logics is more naturally
viewed as one of negation is not, as far as I can see, a matter of significant
philosophical interest. But it is at least telling about how people have thought
about the difference between the two logics and also why negation has been a
recurring problem in manifest forms for constructivism. In what follows I wish
to illustrate the significance of these problems for constructivism generally and
for more narrow programs which may be broadly construed as constructivist
in spirit.

---

[8]A semantics, or the class of structures with respect to which the semantics is given,
characterizes a logic if the logic is sound and complete with respect to that semantics.

Problems with intuitionistic negation

According to the standard intuitionistic (i.e. BHK) interpretation, a negation $\neg A$, defined as $A \to \bot$, is true just in case there is a procedure transforming any proof of A into a proof of an absurdity $\bot$. Now supposing that A is (logically or necessarily) false, there is no proof of A in which case it is vacuously true that there is a procedure transforming any proof of A into a proof of $\bot$. Indeed *every* procedure would qualify as one transforming any proof of a falsehood into a proof of anything.

However this interpretation of negation may seem unsatisfying for a number of reasons. For one, a common intuitionistic interpretation regards a proposition as a particular kind of construction such as the canonical construction which is its proof under some given notion of canonicity according to which a canonical proof is unique (modulo some congruence relation). Now if propositions just are canonical constructions then no sense can be made of a contradictory or absurd proposition required in the understanding of negation as implication to absurdity. For the purported proposition expressed by $\bot$ would not exist since there is no proof, let alone a canonical one, of $\bot$, i.e. $\bot$ fails to express a proposition and is hence meaningless. An objection of this form, not against negation directly but against contradictories more generally, was raised by Hans Freudenthal in [Fre36].

This is admittedly not the only substantive intuitionistic interpretation of a proposition and those that have become more favorable after the latter part of the twentieth century take a proposition to be a *possibly null* collection, e.g. of proofs (the Curry-Feys-Howard interpretation), of solutions to e.g. word problems (the Kolmogorov interpretation), or of methods of fulfilling intensions (the Heyting 1931 interpretation). On these more favorable interpretations there is nothing problematic about *the* absurd proposition since it is identified with the null class and so there is, at least in this respect, nothing problematic about negation defined as implication to absurdity.

However, the problems with interpreting negation intuitionistically do not stop here. The meaning of absurdity has been and still is a problem for inferentialists. If the meaning of a connective is bestowed on it by some set of inferential rules, e.g. its introduction and elimination rules in some natural deduction system satisfying certain meaning-theoretic constraints such as harmony, then what meaning, if any, is attached to $\bot$? There are two main answers to this question.

To avoid vicious circularity we cannot give the meaning of negation in terms of absurdity and conversely give the meaning of absurdity in terms of negation. This makes the obvious candidate for an introduction rule for absurdity, viz.

$$\frac{A \qquad \neg A}{\bot} \ \bot\text{I}$$

a non-starter when the introduction rule for negation is

$$\begin{array}{c} [A] \\ \vdots \\ \dfrac{\bot}{\neg A} \ (\neg\text{I}) \end{array}$$

since then the meaning of each would depend on the other. We should be able to provide an introduction rule for absurdity that does not involve in an essential way any other connective.

Notice that there is no problem concerning an elimination rule for absurdity and that the following standard elimination rule for $\bot$, viz.

$$\frac{\bot}{A}\ (\mathrm{EFQ})$$

proves unproblematic from an inferentialist point of view at least when taken in conjunction with the usual rules associated to the other connectives where circularity does not arise.

Dummett denies that there is any problem of meaning-circularity concerning negation and that $\bot$ may be given harmonious rules by taking EFQ as its elimination rule and the following as its introduction rule:

$$\frac{p_1 \qquad p_2 \qquad \cdots}{\bot}\ \bot\mathrm{I}$$

where the p_i exhaust the class of atomic sentences. Indeed he says "The constant sentence $\bot$ is no more problematic than the universal quantifier: it is simply the conjunction of all atomic sentences" [Dum91, p. 295]. But clearly it is more problematic. For one, it is *infinitary* assuming, as it is usually assumed, that the language contains infinitely many atomic sentences. We cannot assume that only a potential infinity is demanded since any finite set of atomic sentences might just as well be consistent. What is *prima facie* required in specifying the rule is that the actual infinity of atoms be taken as premises unless there are independent grounds for holding that there are only finitely many atomic sentences.

There are grounds for holding that there are only finitely many atomic sentences but Dummett himself never held this. For instance one might promote the atomic sentences to having a special status which implies there being only finitely many of them. One way to do so is to defend atomic sentences as expressing fundamental facts such that a complete description of "reality" may be given by the conjunction of all atomic sentences, a view central to the atomism of Wittgenstein. But, again, Dummett did not construe them this way and, moreover, this response to the infinitary objection invokes a lot of contentious semantic and metaphysical machinery that should not complicate an introduction rule for $\bot$ especially if our understanding of $\bot$ is supposed to be manifestable.

But more importantly, even assuming its being infinitary is not a problem, there is the problem of its being *dynamic*. If the language under consideration, i.e. the one for which Dummett's proposed $\bot\mathrm{I}$ is to be sound, is a natural language, then the rule must be construed as being *open-ended* in the sense that its soundness is preserved under both extensions *and contractions* of the language. For natural language expressions both come and go. The problem is that, while soundness may be preserved under extensions, it is certainly not under contractions, a point which Dummett acknowledges but fails to appreciate. While it is unlikely that English will ever contract to the point where its atomic sentences, if we are able to distinguish such a class in the first place, form an intuitionistically consistent set, the possibility surely cannot be ruled out on *a priori* grounds, and that is enough to undermine Dummett's proposed $\bot\mathrm{I}$.

In other words, Dummett's proposed ⊥I rule is language-dependent in a serious way.[9] There are infinitely many languages for which the conjunction of all atomic sentences is consistent and so should not yield ⊥. Indeed most logical languages are like this. In richer languages such as that of Heyting arithmetic this is obviously not the case since even atoms by themselves, such as $0 = 1$, may be inconsistent. As one of Dummett's programs was to extend verificationism (*à la* intuitionism) in mathematics to verificationism in natural languages such as English, it might only be fair to read ⊥I as being given relative to English or appropriately similar languages in which case the set of atomic sentences, at least if those are delineated in a straightforwardly syntactic fashion, will be inconsistent. But then Dummett needs an argument to the effect that his proposed ⊥I is indeed open-ended.

But Dummett's way is just one as an answer to the the question of whether ⊥ has meaning according to the inferentialist. Another answer is given by Prawitz who holds either that the introduction rule for ⊥ is the "empty rule" or that there is no introduction rule for ⊥. These answers are not equivalent. An inferentialist who holds that only the introduction rules confer meaning on expressions may have no problem with ⊥'s lacking an introduction rule, but an inferentialist who thinks that both the introduction *and elimination* rules together confer meaning on an expression can only opt for the former answer, that the introduction rule for ⊥ is the empty rule. While this likely will not be problematic for modern inferentialists, it will be for those sympathetic to Griss and his rejection of the existence of the null class, the extension of the empty rule if we construe functions class-theoretically. The null class was the source of considerable worry before modern set theory accepted its existence as following from the usual set or class theories such as **ZF**.[10]

Negationless mathematics

On liberal forms of constructivism an object or property exists only if there is a possible construction of it, and on more conservative forms the object or property has to have been actually constructed, i.e. a merely possible construction will not do. Suppose meaning is constrained by existence in the sense that a sentence A has meaning only if every property and object mentioned in A exists. We then have to reject as meaningful what is ordinarily taken to be perfectly meaningful. For instance, the sentence

(*) There are no square circles

will have no meaning since the property of being a square circle cannot be constructed and, hence, cannot exist.

More generally suppose that $A(t)$ is a false mathematical statement. Since $A(t)$ is mathematical, it is necessarily false. Thus not only is there no construction of the property $\lambda x.\langle A(t)\rangle$ expressed by $A(t)$, there could be no construction of the property and so $A(t)$ is meaningless. In other words no false mathematical statement is meaningful.

---

[9]He says that negation lacks the 'invariance' property. See [Dum91, p. 296] for more discussion.

[10]An interesting discussion concerning the history of the reluctance to grant the existence of the null class (e.g. by Russell and Dedekind) can be found in [Kan03].

The view just sketched was first propounded by Griss[11], mentioned already in section 1.1, who held essentially all of Brouwer's views regarding intuitionistic mathematics except for one important one, viz. that false or refutable sentences are meaningful and that the negation of a statement can be established by deducing from the statement an absurdity. For Griss only true sentences are meaningful. He gave a negationless reconstruction of a reasonable fragment of intuitionistic mathematics. In fact some proposed first-order constructive negationless theories of arithmetic have been shown intertranslatable with Heyting arithmetic.[12]

Let be # the "positive" concept of *apartness* which, in the context of negationless theories, is used as a substitute for inequality, $\neq$.[13] For instance, a negationless recasting of (*) is

(**) "If S is a square and P any point, then we can find points Q and R on the boundary of S, such that $PQ \# PR$"[14]

A simpler example is the positive recasting of $x \neq y$ in the theory of strict linear orders by the statement $x < y \lor y < x$. Indeed the relation just so defined, viz. $x < y \lor y < x$, is an apartness relation.

One apparent oddity of negationless systems from a semantical point of view is that the class of meaningful propositions is not closed under conjunction. For even if 'x is square' and 'x is circular' are meaningful, their conjunction is not since there is no way to construct a square circle. Another way of putting this is to say that in the negationless theory of classes where the null class fails to exist, the intersection of two classes is not guaranteed to yield another class. The problem of developing syntactical systems adequate to the intended semantics of negationless theories has proved notoriously difficult.

Strong negation

Where does this leave negation if the negation $\neg A$ of a sentence A is true iff A is false and all false statements are meaningless? Nelson's [Nel49] may be regarded as having generalized the notion of apartness as a positive substitute for inequality to a positive notion of constructible negation. The generalization, then, was from a positive notion of negation applied only to equality formulae to that same *positive and constructible* notion of negation applied to arbitrary formulae. In the theory of linear orders to say that x is apart from y is to say that there is a method which tells us whether $x < y$ or $y < x$. The method gives us a procedure for determining the *falsity* of the identity of x with y without recourse to absurdity. Likewise there exist positive procedures

---

[11]See e.g. [Gri46].

[12]See [Vic00a] and [Vic00b]. It's not clear that translational equivalence in the foregoing sense carries much currency. Certainly it does not by itself show that two theories are equivalent as regards what they are *intuitively* able to express.

[13]Apartness is used also outside of negationless theories of mathematics where $x \# y$ is strictly stronger than $x \neq y$. On apartness, see [Hey71, §2.2.3] and [vD08, pp. 175–177].

[14]This example is taken from [Hey71, p. 124] but he works through the contradiction only antecedently in [Hey55, p. 92]: "[i]n this example it is easy to deduce the contradiction, for let S be a square and P any point, a one of the sides of the square, Q the foot of the perpendicular from P to a, R a point on a different from Q, then the distances PQ and PR are unequal; but if S were a circle, we could take its centre for P and PQ would be equal to PR. The contradiction is found".

for determining the falsity of other kinds of statements and whenever we have such a procedure, say for determining the falsity of A, we should be able to assert $\neg A$ in the constructivist spirit envisaged by Griss. For what $\neg A$ asserts in such a case is not the meaningless $A \to \bot$ but that e.g. there exist, in the intuitionistic sense, numbers realizing the falsity of A.[15]

In Kripke semantics for intuitionistic propositional logic **IPC**, strong negation $\rightharpoonup$ can be modeled by distinguishing verification at a point in the model from falsification at a point. Let $M = (W, \leq, V^+, V^-)$ be a Kripke model with W a non-empty set, $\leq$ a preorder on W and V^+ and V^- valuations assigning upsets of W to atoms. X is an upset of W iff $b \in X$ whenever $a \in X$ and $a \leq b$. Valuations must satisfy the restriction that for each atomic p, $V^+(p) \cap V^-(p) = \emptyset$; in other words, atomic sentences are never both verfied and falsified. This restriction holds for arbitrary formulae given the verification and falsification conditions listed below, a proof of which is omitted here.

The logical vocabulary of the language is given by the following verification and falsification conditions. We write $M, a \models^+ A$ for 'A is verified under M at a' and $M, a \models^- A$ for 'A is falsified under M at a' where $a \in W$. This gives us the following (suppressing M):

- $a \models^+ p$ iff $a \in V^+(p)$

- $a \models^- p$ iff $a \in V^-(p)$

- $a \models^+ \bot$ never holds

- $a \models^- \bot$ always holds

- $a \models^+ A \wedge B$ iff $a \models^+ A$ and $a \models^+ B$

- $a \models^- A \wedge B$ iff $a \models^- A$ or $a \models^- B$

- $a \models^+ A \vee B$ iff $a \models^+ A$ or $a \models^+ B$

- $a \models^- A \vee B$ iff $a \models^- A$ and $a \models^- B$

- $a \models^+ A \to B$ iff $\forall b \geq a$, if $a \models^+ A$ then $a \models^+ B$

- $a \models^- A \to B$ iff $a \models^+ A$ and $a \models^- B$

- $a \models^+ \rightharpoonup A$ iff $a \models^- A$

- $a \models^- \rightharpoonup A$ iff $a \models^+ A$

An argument (Γ, A) is valid on the semantics, written $\Gamma \models A$, iff for all models M and a of M, if $M, a \models^+ B$ for all $B \in \Gamma$ then $M, a \models^+ A$. If (Γ, A) is a valid argument we write $\Gamma \models A$.

An axiomatization of the class of arguments valid on the semantics is got by taking an axiomatization of **IPC** and adding to it the following axioms:

1. $\rightharpoonup A \to (A \to B)$

2. $\rightharpoonup(A \to B) \leftrightarrow (A \wedge \rightharpoonup B)$

---

[15]The sense of realizability is that of [Kle45] and is employed in the interpretation of constructible negation in [Nel49]. Strictly speaking, then, the constructive interpretation given to the language in [Nel49] diverges from the intuitionistic interpretation.

3. $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$

4. $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$

5. $\neg\neg A \leftrightarrow A$

6. $\neg\neg A \leftrightarrow A$

It should be noted that the logic thereby obtained is not closed under substitution of provable equivalents only when the formulae involved contain occurrences of strong negation.

Besides providing a strongly constructivist interpretation to negation, strong negation has been put to other uses. It has been, for example, used in solutions to the knowability (or Fitch's) paradox according to which the verificationist principle,

(VP) $A \rightarrow \Diamond KA$

implies what might be taken to be the implausible claim that no truth is unknown, i.e. that $\neg(A \wedge \neg KA)$ (assuming A is taken schematically). In VP, $\Diamond$ is a possibility modal closed under provable implications (if $A \vdash B$ then $\Diamond A \vdash \Diamond B$) and $\vdash \neg \Diamond \bot$ for $\bot$ an arbitrary **IPC**-contradiction, and K is a knowledge operator satisfying distribution over conjunction ($K(A \wedge B) \vdash KA \wedge KB$) and factivity ($KA \vdash A$). A proof runs as follows:

1. $A \wedge \neg KA$ hypothesis
2. $\Diamond K(A \wedge \neg KA)$ 1,VP by $\rightarrow$E (modus ponens)
3. $\Diamond(KA \wedge K\neg KA)$ 2 by K-distribution and closure of $\Diamond$ under $\vdash$
4. $\Diamond(KA \wedge \neg KA)$ 3 by factivity and closure of $\Diamond$ under $\vdash$
5. $\neg(A \wedge \neg KA)$ 1,4 by $\neg$I and $\vdash \neg \Diamond \bot$

Dummett [Dum09] finds nothing wrong with the "implausible" conclusion taken on an intuitionistically acceptable reading of negation (on which more later). In the original formulation of the paradox by Fitch (see [Fre63]), the conclusion was the stronger $A \rightarrow KA$ which classically but not intuitionistically follows from $\neg(A \wedge \neg KA)$. Intuitionstically we still have the contrapositive $\neg KA \rightarrow \neg A$ following from $\neg(A \wedge \neg KA)$ which can look at least as bad as, and if not worse, than $\neg(A \wedge \neg KA)$.[16] Since the conclusion $A \rightarrow KA$ is not available intuitionistically and since verificationism may be more plausibly held in an intuitionistic light, for present purposes it is best to ignore the less plausible original "classical" formulation of the paradox.

One way to block the foregoing derivation ('the derivation' for short) is to give up $\neg$I. Since strong negation fails to satisfy $\neg$I one might think that a possible solution to the paradox is to employ strong negation in place of intuitionistic negation in the derivation. This solution is proposed by Wansing [Wan02] but he notices that there is still an additional worry. Even if

(1.1) $\neg \Diamond K(A \wedge \neg KA)$

---

[16]The contrapositive looks worse since it is not the negation of a Moorean sentence, and negations of Moorean sentences strike us as entirely assertible. Of course what is not assertible is the negation of the existential $\exists A(A \wedge \neg KA)$. A fair formalization of the paradox should make the propositional quantifiers explicit.

is postulated as an additional hypothesis, where $\neg$ denotes strong negation[17], since then

$$(1.2) \qquad (A \wedge \neg KA) \rightarrow (\Diamond K(A \wedge \neg KA) \wedge \neg \Diamond K(A \wedge \neg KA))$$

would be derivable. However, it is not clear this is really a problem since one is never in a position to assert (any instance of) the antecedent of (1.2) (definitely if the knowledge norm for assertion holds and given the assumption (1.1)), being (an instance of) a Moorean sentence such as 'It's raining but I don't know it'. In fact, if $\neg$ is strong negation then $A \wedge \neg KA$ would be for a verificationist holding (1.1) something of a contradiction in which case it should come to no surprise to her that (1.2) hold, a theorem reminiscent of EFQ.[18]

In any case, there is another problem with this strategy that is independent of whether or not we accept (1.1) and the assertibility of Moorean sentences. Not every instance of 'It is not known that A' can be given a faithful translation using just strong or intuitionistic negation (and the other positive intuitionistic connectives) that is due to the following two considerations. The first is that, if the derivation is formalized using intuitionistic or stronger negation, Dummett's response to the paradox in [Dum09] stands. For $\neg(A \wedge \neg KA)$, the supposedly abhorrent conclusion following intuitionistically from VP, would assert that "It is in principle impossible for us to be in a position to assert both that A and that it is impossible for us to be in a position to assert KA" which is entirely reasonable for a verificationist to hold. (In the case of strong negation, 'impossible' in Dummett's quote would be given an even stronger, though intuitionistically palatable, reading.)

The second is that both strong and intuitionistic negation are *generally* too strong. The claim that the proposition that A is unknown *at present* will, in many cases, be weaker than the claim that an assumption of KA leads intuitionistically to an absurdity. For there will be cases in which A is verifiable and unknown even though it is not *definitely* or otherwise *refutably* unknown, each of which is asserted by the respective intuitionistic negation $\neg A$ of A and its strong negation $\rightarrow A$.

While strong negation might have nice proof-theoretic properties for warding off the knowability paradox, its intended interpretation does not *always* provide a faithful translation of sentences of the form 'It is unknown that A'. (In interesting cases it does. If we read $\rightarrow A$ as 'It is definitely not the case that A', where 'not' has a constructive flavor, then it is reasonable to formalize 'Goldbach's conjecture is unknown' using strong negation, for we definitely know that the conjecture has not yet been proved and is hence unknown.) For an assertion of the strong negation of A is an assertion that there is a procedure which provides suitable evidence (in the case of arithmetic, natural numbers) demonstrating the falsity of A. For instance if A is 'Everything is not red' then the procedure would yield red objects, or if one thinks evidence is propositional, the proposition describing the existence of any such object.

One might also think that if a statement or proposition "says" that A and that A logically implies that B, then the statement also, at least implicitly, says

---

[17]Recall that (1.1) is intuitionistically derivable by just the properties of $\Diamond$ and K previously mentioned together with $\neg$I which does not hold for strong negation.

[18]While I do not think an acceptance of (1.2) is problematic, Wansing does. To overcome the problem, he endorses a relevant implication on which the problematic (1.2) is not derivable.

that B. If I assert that 3 is prime and odd then I have asserted that it is odd. But then we have another easy argument against formulating the knowability derivation using strong negation on the supposition that intuitionistic negation is already too strong in general to formulate the derivation since $\rightarrow A \vdash \neg A$ (though this implication fails in the logic **N4** employed in [Wan02] and hence this easy argument does not apply there).

One might not care for the intended semantics of strong negation, opting instead to care only about what proof-theoretic properties it has. But then there are plenty of equally good ways out of the paradox: just select your favorite negation that fails to possess the properties required for the derivation to go through. The problem with this sort of "solution" to the paradox is that one does not want to merely block the derivation, one wants to know why the alleged negation appealed to in a solution of this sort should be taken seriously from a verificationist point of view. The point, to reemphasize, is that any good solution to the paradox that blocks the derivation via negation should be such that the intended semantics of the negation gives a faithful translation of the ordinary language instances involved in the derivation. While intuitionistic and strong negation work fine for certain instances of A of (certain instances of) the derivation, they are not suitable in general, e.g. when it has not been determined whether or not A is in fact known).

Now suppose we formulate the paradox using just an empirical negation construed broadly along intuitionistic lines. Then it is far from clear that the derivation would go through. Indeed, if empirical negation has the properties proposed in chapter 3 then the derivation does not go through and, moreover, it is argued that the proposed negation provides a faithful formalization of the derivation in its most general setting.

Empirical negation

I argued that the negation employed in the knowability paradox can be neither intuitionistic nor strong negation if (i) in agreement with [Dum09], the purported air of paradoxicality is to be genuine, and (ii) sentences of the form 'It is unknown that A' are to be faithfully regimented into a language acceptable to a constructivist or verificationist given the intended meaning of the respective negations (extended in some naïve way to empirical discourse). A faithful translation of such sentences requires a weak empirical negation $\sim$ according to which $\sim A$ says merely that A is unwarranted *at present*. But there are additional related and pressing reasons for providing an adequate analysis of empirical negation that are brought to light by Williamson in [Wil94].

In [Wil94], Williamson poses a problem for verificationism by arguing that statements of the form

(*) A may never be decided

are intuitionistically inconsistent on some suitable extension of intuitionistic semantics to empirical discourse. (Notice that statements of the form (*) are empirical, even when A is mathematical.) For example the statement 'For all we know, Goldbach's conjecture will never be decided' turns out intuitionistically equivalent to '$1 = 0$'. This is quite obviously unpalatable to any sort of constructivist.

Williamson's argument is somewhat overstated since there are statements that are *provably*, even by constructive means, undecidable. It follows that it is intuitionistically consistent to say that such sentences may never be decided. For let A (e.g. a Gödel sentence) be such a sentence. Then we have, and hence *know*, a proof of the undecidability of A, i.e. that neither A nor its negation is provable. Assuming that once we know a proof of a statement this knowledge persists indefinitely into the future, it follows that 'A will never be decided' is consistent with what we know. But that is just to say 'A may never be decided' is intuitionistically consistent.[19] Williamson's argument must then be restricted to sentences that are in fact *presently* undecided, such as Goldbach's conjecture.[20]

Williamson offers a simplification of his argument according to the following reasoning. If we can show the inconsistency of a non-epistemic claim 'A will never be decided', then we know it to be inconsistent, and so the weaker epistemic claim 'A may never be decided' must also be inconsistent. Moreover the converse is obviously true: the inconsistency of the weaker claim implies the inconsistency of the stronger one. This allows us to simplify the consistency question by ignoring the epistemic noise introduced by the (*) version, so that we may focus solely on its non-epistemic variant.

I do not think Williamson's argument is convincing for some of the reasons I give below, but I do think it is instructive for thinking about some of the problems associated with intuitionistic negation and empirical discourse. One problem is that there is no way to express that a statement *lacks warrant*. The closest the intuitionist can get is to express that any warrant for asserting A is just as much a warrant for asserting an absurdity, and clearly this is much stronger than simply claiming that A lacks warrant. Now how does this relate to Williamson's argument? The statement that A will never be decided just is the statement that it will always be the case that A *lacks warrant* which the intuitionist is at a loss to express without recourse to empirical negation.

Williamson's argument

The following is Williamson's argument. Assume intuitionistic semantics has been extended to empirical discourse, and let KA mean 'at some past, present or future time someone possesses a warrant to assert A'. Then a statement of the form

(**) A may never be decided

may be formalized as $\neg(KA \lor K\neg A)$ which is intuitionistically equivalent to $\neg KA \land \neg K\neg A$. But $\neg KA$ implies $\neg A$ (shown in a moment) which, together with $\neg KA \land \neg K\neg A$, yields $\neg A \land \neg\neg A$. For suppose $A \land \neg KA$ could be warranted (at some time), i.e. $K(A \land \neg KA)$. As warrant distributes over conjunction, $KA \land K\neg KA$. By the "factivity" of K we obtain $KA \land \neg KA$ and so $\neg(A \land \neg KA)$. It follows intuitionistically that $\neg KA \to \neg A$. But then $(\neg KA \land \neg K\neg A)$, expressing that A will (or equivalently, may) never be decided, implies $\neg A \land \neg\neg A$,

---

[19] Non-temporal modalities are always to be read as epistemic.

[20] One may object to the overstatement point here by holding that the intuitionistic notion of proof is *informal* and not *proof in* **L** for any given formal system **L** in which case it is arguable that there in fact is a proof, *in the informal sense*, that arithmetic, *taken as an informal theory*, is incomplete. The point is well-taken.

an intuitionistic contradiction. We may thus conclude that sentences having the form (**) are intuitionistically inconsistent. A more detailed discussion of these points is found in chapter 3, section 3.7.

Expanding K to explicitly reveal its temporal nature will not help the intuitionist. If we introduce temporal operators P ('it was the case that') and F ('it will be the case that') into our language, then KA is equivalent to $PA \lor A \lor FA$ on Williamson's reading of A as 'A is warranted (at present)', and so $\neg(KA \lor K\neg A)$ is equivalent to $\neg((PA \lor A \lor FA) \lor (P\neg A \lor \neg A \lor F\neg A)$. But this too is intuitionistically inconsistent on straightforward extensions of intuitionistic logic to include temporal operators having properties validated on either linear or branching time semantics.[21]

Williamson suggests that one way out of the problem is to formulate undecidability claims of the (**) form using strong negation. He rejects the move given a particular reading of K according to which KA and A have the same verification, but different falsification, conditions. This is itself quite puzzling.[22] But for reasons mentioned earlier, we saw that both strong and intuitionistic negation are too strong in any case to act as the sort of empirical negation required in making claims of the form (**). Thus a weak empirical negation provides a more plausible solution to Williamson's objection that claims of the form (*) are intuitionistically inconsistent.

We see there are a number of convincing reasons for the intuitionist to be able to express empirical negation. The semantics and logic for such a negation is given and motivated in chapter 3 and the logic is proved sound and complete with respect to the semantics.

1.4   Consequence relations for speech acts and propositional attitudes

The speech act of denial and the propositional attitude of rejection have gained renewed importance in the philosophy of logic. One reason for this comes from the literature on truth. One initial reaction to the liar paradox is to say that the liar sentence,

(λ) This sentence (i.e. λ) is false,

is gappy, that is, neither true nor false. The solution works just fine for this particular case of the liar but when we consider the strengthened liar,

(λ^*) This sentence (i.e. λ^*) is not true,

we land right back in paradox, assuming the usual properties of truth, namely that

$$\frac{A}{T^\ulcorner A^\urcorner} \, (T\mathrm{I}) \qquad\qquad \frac{T^\ulcorner A^\urcorner}{A} \, (T\mathrm{E})$$

---

[21]One way to see this is to show that K, as we have abbreviated it in terms of P and F, still distributes over conjunction, which is all Williamson's argument requires.

[22]His discussion in the earlier part of [Wil94] strongly indicates that the verification conditions of the two are distinct with the verification conditions for A being strictly more difficult to satisfy than those for KA. For if the two have equivalent verification conditions then in the context of his earlier discussion, the formalization of (**) as $\neg KA \land \neg K\neg A$ would turn out as essentially an abbreviation for $\neg A \land \neg\neg A$ which is unsatisfactory as a fair intuitionistic formalization of the claim.

hold unrestrictedly (i.e. for arbitrary A), where $\ulcorner A \urcorner$ is a quote name or Gödel number for A and consequence is classical.[23] To see this, assume $T\ulcorner \lambda^* \urcorner$. By TE we have $\neg T\ulcorner \lambda^* \urcorner$, whence $\neg T\ulcorner \lambda^* \urcorner$ follows by $\neg$I from no assumptions. It follows by TI and the definition of λ^* that $T\ulcorner \lambda^* \urcorner$. By $\neg$I and double negation elimination we obtain $T\ulcorner \lambda^* \urcorner$ from no assumptions. Assuming any logic for which EFQ holds in the form $A, \neg A \vdash B$, triviality ensues.

The gap theorist is still open to holding that λ^* is gappy but if she does so she will have to adjust either her logic or else her truth principles accordingly in order to block triviality. One popular approach is to take up the former by weakening one's logic, for example, by adopting a gappy logic, a move that seems natural for a gap theorist to make. However, this sort of approach, of which Kripke's is a prime example, is open to a *prima facie* difficult challenge. For since the gap theorist admits that λ^* is neither true nor false, what λ^* says *is the case*. The problem is that 'is the case' must mean something different from 'is true' and it is difficult to see what that might be. One reading is that it means 'is assertible'. But then the gap theorist must hold that λ^* is assertible while *knowing* at the same time that it is not true. Such a concession flies in the face of the usual *non-pragmatic* conception of assertibility according to which A is unassertible by S if it is regarded by S to be untrue, a condition that is considerably weak as far as constraints on assertibility and unassertibility go. (Notice that this is *not* equivalent to what might at first glance appear to be its contrapositive, viz. "If A is regarded by S to be true, then A is assertible". Indeed this would-be contrapositive is implausibly strong though not clearly on a non-pragmatic view of assertibility according to which that notion is constrained by rationality rather than pragmatic factors. For that A is assertible does not imply that it *ought to be* asserted.)

The gap theorist typically responds to this challenge by claiming that her utterance of λ^* is *not* an assertion of that sentence (or the proposition expressed by that sentence, supposing that assertions act on propositions rather than sentences, but more on this just below), rather it is a *denial* of that sentence (or, again, the proposition expressed by the sentence). The gap theorist now has to take each assertion and denial as primitive, where the "Fregean" needs only one (and indeed, it does not matter which though Frege took assertion as the primitive speech act) by holding that a denial that A is an assertion that $\neg A$.[24] In other words, the gap theorist has to deny that there is a close connection between negation and denial.

Now positing two speech acts as primitive instead of just one of them is not necessarily a problem provided doing so is grounded in reasons independent of merely addressing the above challenge. Otherwise the positing of the additional primitive looks *ad hoc*; there has to be some worthy additional payoff for the theoretical cost of positing the additional primitive. The gap theorist can avoid *ad hocery* by endorsing the very same theory of assertion and denial the Fregean endorses under some intuitive notion of *sameness*. For instance, each can hold that

---

[23] The rules are sometimes labeled T-IN and T-OUT respectively, but I see no reason to break standard rule-naming conventions in this particular case even if an application of TE results in a conclusion containing more occurrences of T than occurs in the premise.

[24] Frege famously held the view that the denial that A just is the assertion that $\neg A$. See [Fre60].

(**Assert**) A is assertible iff A is true;

(**Deny**) A is deniable iff A is untrue.

But since a sentence is not untrue iff its negation is true according to the gap theorist, Deny can no longer be seen as a special instance of Assert and this explains why the two need to be taken as primitive for the gap theorist.

However, Deny is extremely implausible for a gap theorist to hold, at least when taken to hold for arbitrary sentences. One reason that a sentence may be untrue is because it is not truth-apt, for example, it might be a command such as 'Take out the garbage'. If denials act on propositional contents which are themselves necessarily truth-apt, then it makes no sense to deny the content of 'Take out the garbage' since whatever that content is, it is not a proposition. Moreover, even if we restrict ourselves to truth-apt sentences, Deny still seems implausible according to some widely-defended gap theories as the following considerations bring to light.

Consider the gap theorist who denies that λ^* is either true or false because it fails to express a proposition. Such a gap theorist cannot deny that λ^* since, again, λ^* fails to express a proposition and denials act on propositions. Where there's no content, there's no denial. Unless denials act on something other than propositions or unless a denial of λ^* is not a straightforward denial of the content of λ^* but the denial of some other proposition, e.g. the proposition that λ^* is true, the gap theorist will have to reject Deny in favor of something like the obvious variant:

(**Deny$'$**) A is deniable iff A is false.

I think this is the most intuitive and plausible position for the gap theorist to take, though I consider the aforementioned options—that is, the idea that a denial of A may involve more than just the content of A—in chapter 4, section 4.2 and find them wanting.[25]

With an endorsement of Deny$'$ comes the need for a distinction between two notions of consequence, viz. consequence taken as (i) the preservation of truth from premises to conclusion and (ii) the preservation of falsity from premises to conclusion. In classical logic each is definable in terms of the latter for if the truth of each member of Γ implies the truth of A then at least one member of Γ is false if A is false. Likewise if the falsity of each member of Γ implies the falsity of A then the truth of A implies the truth of at least one member of Γ. In other words, classical logical truth preservational consequence anti-preserves falsity and its converse preserves falsity. However in a gappy logic (more precisely a gappy semantics, though we may think of logics as coming with a semantics, and we shall do so sometimes for convenience) each of the two notions of consequence is not definable in terms of the other. For consequence as truth preservation anti-preserves *untruth*, not falsity, the notion figuring centrally in Deny$'$.

As it turns out there are a variety of speech acts and propositional attitudes other than assertion, acceptance, denial and rejection that may be characterized or rationally constrained by various notions of consequence distinct from

---

[25]One standard objection to thinking that the denial of λ^* is the denial of the content of $T\lambda^*$ is that the attribution of different properties of λ^* are being denied in the respective denials of λ^* and $T\lambda^*$—in the former it is untruth and in the latter it is truth. They must therefore be different acts.

consequence as truth preservation. One of those is doubt or dubitability. Consider a three-valued semantics where the values correspond to truth, falsity and some notion of *indeterminacy*. Then one might think the following characterization of dubitability ought to hold:

(Doubt) A is dubitable iff A is indeterminate

where A may be restricted to truth-apt sentences. This suggests the following principle concerning the rationality of doubting:

(DC) If each member of Γ is indeterminate implies that A is indeterminate then A is dubitable if each member of Γ is.

Clearly the principle can be formulated only if we have at our disposal a notion of consequence as indeterminacy preservation. More importantly, we cannot hold people hostage to the principle if there are no rules they can follow which take them from indeterminate sentences (or whatever the truthbearers are, primarily or derivatively) to other indeterminate sentences. We should be able to hold someone as being rationally irresponsible for doubting each member of Γ on the grounds of their indeterminacy while believing A if the indeterminacy of each member of Γ implies the indeterminacy of A just as we would hold them account if they believed each member of Γ on the grounds of their being true while rejecting A even though Γ implies A in the truth-preservational sense.

With a marked separation of negation and falsity comes a need for distinguishing various notions of consequence and the role they play in rationally constraining various speech acts and attitudes. Various of these relationships are explored in chapter 4 with a focus on consequence as indeterminacy preservation and its relation to the attitude of dubitability. A natural deduction system for indeterminacy and falsity preservation over Kleene's matrix K_3 is given and proved sound and complete with respect to the class of K_3 matrices.

1.5 Negation, truth and falsity

In [Fre48], Frege argued that truth values are *sui generis* objects. He called one *the true* and the other *the false*. Besides their being regarded as referents of sentences, he said little more about their nature. The only thing he said in this regard, which is too vague to be of much interest, is that "by the truth value of a sentence I understand the circumstance that it is true or false" [Fre48, p. 216]. According to Church (*Introduction to Mathematical Logic*, p. 25, 1956), the only person before Frege to have explicitly mentioned truth values, *taken in its modern sense*, was C. S. Peirce in 1885 in a paper *On the Algebra of Logic: A Contribution to the Philosophy of Notation*. It is duly noted in [SW10], however, that 'truth value' was used nine years earlier in 1882 by Windelband in his paper *What is Philosophy?*, though he did not mean by it the reference of a sentence or sentential function.

Over one hundred years later and we find investigation into the nature of truth values regarded as objects rather than properties to be rather sparse. (Of course one might hold that properties themselves are objects, in which case by 'object' understand me to mean "first-order" object.) This is surprising given the intimate connection between truth, falsity and negation and the fact that each of these three notions has been *intensely* investigated over those hundred

and some odd years. It is nearly standard, and not just in a classical setting, to define negation as a toggle between truth and falsity (an account called 'TC' in chapter 2 to which the reader is referred for details): a negation $\neg A$ is true iff A is false. So there is a close link between falsity and negation or contradictories; a negation being true in virtue of some corresponding sentence being false. But we may also turn the dependence on its head since it is also standard affair, and not just in a classical setting, to define falsity in terms of truth of negation. Concerning the latter reduction of falsity to truth and negation, however, such a reduction must be accompanied by an account of the truth of negations, the so-called "negative truths", without any vicious recourse to falsity, explicit or otherwise, if it is to be even minimally adequate.

Providing a theory of negative truths has proved notoriously difficult. Indeed, an upholding of the distinction between "positive" and "negative" truthbearers (I'll use 'statement' to refer to whatever entity the primary truthbearers are) has proved difficult. (A syntactic distinction is unproblematic but it also bears no philosophical weight.) A renewed interest in correspondence theory in the modern form of truthmaker theory has brought renewed interest in providing purely negative-free accounts—by which I mean accounts that do not posit negative entities such as totalities—of truthmakers for negations. The problem is that none of these accounts, barring those of the atomists, are particularly compelling from an ontological point of view and even atomism appears to be too strong. It is all too easy positing the ontological atrocities that are negative entities in order to endorse a theoretically driven principle such as truthmaker maximalism. The alleged payoff just isn't worth it.

Of course, one need not posit, like Russell, negative facts nor, like Wittgenstein, only facts corresponding to atomic truths. One way out of the tension Russell felt regarding negative facts, the seeds of which are already present in [Rus10], is to posit just positive facts but two *ways* of corresponding to those facts. Russell appears to suggest this in the following but he intends something quite different since he feels the pressure to ultimately posit negative facts:

> The essence of a proposition is that it can correspond in two ways with a fact, in what one may call the true way or the false way...Supposing you have the proposition "Socrates is mortal", either there would be the fact that Socrates is mortal or there would be the fact that Socrates is not mortal. In the one case it corresponds in a way that makes the proposition true, in the other case in a way that makes the proposition false. That is one way in which a proposition differs from a name. [Rus10, p. 38]

However, while it appears we have eliminated the need for negative facts, we have simply traded one problem for another, for now we have to explain what the difference between the two ways of correspondence is, and it is highly dubious that an account more satisfactory than one which posits a single way of correspondence and two types of facts such as Russell's is forthcoming.

So the question is: Can a plausible and ontologically modest reduction of falsity to truth of negation be given? I think it certainly can. I also think that such an account can be useful for other reasons. One of those is by illuminating the nature of truth-value-as-object. If we are reducing falsity to truth and negation then it should turn out that falsity-as-object does not exist,

at least not as a *sui generis* object, just as the mental does not exist as a *sui generis* substance if it is reducible in the relevant sense to the physical. Perhaps it is best to say that, here, we have an *elimination* rather than a reduction of truth.[26]

Now think about what this says for semantical theories representationally construed. From the looks of it we should only have one truth value, viz. truth. But the intended semantics of most logics come equipped with more than one value. What should we say, then, about all the values distinct from truth? What exactly can they be representing if there is nothing in reality to which they correspond? The simple answer: nothing.

This looks like a problem but it is not. Every logic for which $A \vdash A$ holds (and even some for which it fails) can be given a one-valued semantics (on which see chapter 5). From a purely formal standpoint, this justifies the elimination of falsity. In fact, it justifies the elimination of any logical value distinct from truth. But this purely formal fact is also justified by the philosophical defense of truth as the maximal truthmaker, a view defended in chapter 5. So justification feeds in both directions, from the formal to the philosophical and conversely. One conclusion to draw from these considerations is that, in terms of priority, negation comes first, falsity—taken as truth of negation—last. This priority was rejected by Russell based upon the confusion that "we may safely treat "false" and "not" as synonyms" ([Rus40, p. 78]), a confusion duly noted in [Ros72].

1.6   The definability of negation as impossibility

To claim that some state of affairs is possible or impossible is a common philosophical affair. It would come as shocking news, then, if philosophers found out that it is impossible to express impossibility! How they might ever come to understand the forgoing sentence, if it even expresses anything meaningful, would appear to go beyond the limits of language. This sounds incredible. But it is not if the language we employ, e.g. English, has the kind of semantics one typically finds in the philosophical literature. By way of analogy, let us first consider definability and expressibility in first-order languages.

It is well known how inadequate first-order languages are at defining certain classes of structures. For example, one can demonstrate a sentence true on only infinite models, but one cannot demonstrate a sentence true on only infinite models of some fixed cardinality. If we allow identity a fixed semantic interpretation, then we get something of the reverse for finiteness: one can demonstrate a sentence true on only finite models of some fixed cardinality, but one cannot demonstrate a sentence true on only finite models generally. The same sorts of considerations carry over to modal languages. The basic modal language, on its usual interpretation, is ill-equipped at defining a number of important frame properties including one central to interpreting negation as

---

[26]As with any "reduction" or "elimination" given via a biconditional, here it goes in both directions: one may "reduce" or "eliminate" truth in terms of falsity and likewise assertion in terms of denial (a point ignored by Frege in his famous reduction of denial to assertion of negation). I will not have anything to say about this here, though I do think there is good reason to favor one reduction over the other not simply because genuine reductions and eliminations are unidirectional and hence there must be something which breaks the symmetry.

impossibility *in the broadest sense*. The reason for this is due to the indefinability of the modality that has come to be called in the literature the "global" or "universal" modality (see [Gor90], [GP92], and [BdRV02]). The global interpretation of necessity has been around a long time. It was reflected in the work of Carnap in the forties, out of which arose a reinvention of the modal system **S5**. His idea was to model a possible world as a maximal set of atomic letters, a "state-description", and to then define necessity as truth in all worlds. An atomic sentence is true in a world (state-description) just in case it is a member of the world and truth for complex formulas is extended from the atomic case in the usual way. If a sentence is not true in every world then its necessitation is false in all of them, an interpretation familiar in the so-called Henle matrices for **S5** (see e.g. [DH01]).

An interesting question arises: What familiar modal languages are capable of defining the global modality and hence "absolute" impossibility? In an attempt to answer this question, I focus attention on just a few languages I find the most interesting. The first concerns conditional logic and propositional quantification, and the second a modality investigated in [GPT87] as a "sufficiency operator", in [Hum83] and [Gor90] as a "complementary" modality, and in [Gol74] and [Dun93] as a negation. There are other interesting ways of defining the global modality which, unfortunately, I have no space to survey here.

The philosophical importance of the global modality

In the usual base modal language (defined in chapter 6, section 6.1) it makes no difference to a modal theory (taken as the set of sentences valid over a class of *frames*) whether the box is interpreted by the universal relation (or equivalently, by no relation) or by an arbitrary equivalence relation. The same is not true when we enrich such an impoverished language. In the propositional case, consider languages with "nominals", a propositional letter true at precisely one world. If @ is a nominal, then $\Diamond@$ is valid over the class of frames with a universal alternative relation, but it is not valid over frames with a mere equivalence relation (or some fixed non-universal equivalence relation). These considerations play an important role also from a philosophical point of view.

The differences between different **S5** interpretations of box are brought out more clearly in the case of theories over classes of *models* rather than frames. The philosophical interest of this is especially highlighted in the setting of first-order modal languages. Consider the oft-discussed Barcan formula (BF),

(BF) $$\Diamond \exists x A(x) \rightarrow \exists x \Diamond A(x)$$

which says, roughly, that all possible things are actual—i.e. that *possibilism*, the thesis that there exist non-actual possibilia, is false. If one gives an interpretation of the modal language according to which $\Diamond$ is not global, then (BF) may be valid over a class of models in which the domain of each alternative of the actual world is a subset of the domain of the actual world, even though there are inaccessible worlds for which this inclusion fails. But then the validity of the Barcan formula does not, after all, imply the falsity of possibilism, for there exist non-actual possibilia at inaccessible worlds. However, if one is in the business of making controversial metaphysical claims, one had better

do so in an appropriately interpreted language, and in this case that means interpreting the modal operators globally. In practice, one simply does this by stipulation—indeed, in such cases one *should* do this by stipulation. In chapter 6 we shall make no stipulations and instead make do with what we have, defining the global modality in terms of other "naturally-occurring" operators.

Let us consider one final property of the global modality which makes it philosophically interesting: viz., being the only (in a certain sense) *non-trivial* modality that is invariant under permutations of (the set of worlds of) Kripke models. (The trivial operator satisfying $A \leftrightarrow \Box A$ is also permutation invariant in this sense, being governed by the identity relation, as is the modality satisfying $\Box A$ for all A, being governed by the empty relation.) Other modal operators, viewed as quantifiers inherently restricted by alternative relations that are not permutation invariant, do not share this property. Why is permutation invariance important? Most significantly, because it is thought (e.g. by Tarski) to provide a criterion of—or perhaps more plausibly either necessary or sufficient conditions for—being a *logical constant*. Under this criterion the **S5** modal operators (box and diamond) turn out to be logical and essentially all others (e.g. all those strictly weaker than **S5**'s) come out non-logical.

So the global modality is philosophically interesting and yet it is not definable in the standard languages one typically encounters. The interest of chapter 6 lies in exploring some languages familiar in the literature that have just enough expressive power for defining negation as impossibility which, taking boolean negation for granted, reduces to the problem of determining which languages have the resources for defining the global modality.

Final remarks

The chapters to follow are essentially self-contained. Reference is made to other chapters but ignoring these would not cause any serious detriment to understanding.

Let us recap. **Chapter 2**, *What is negation?*, discusses the problem of characterizing negation and why it is important to projects such as logical revisionism. **Chapter 3**, *Constructivism and empirical negation*, discusses the problem of empirical negation for intuitionism and provides an account of empirical negation amenable to the intuitionist. **Chapter 4**, *Consequence relations for speech acts and propositional attitudes*, highlights the importance of nonstandard notions of consequence, such as consequence as indeterminacy preservation, required in formulating rationality constraints on various speech acts and attitudes such as denial and doubt. **Chapter 5**, *Truthmaking, negative truths and truth values*, provides both an ontological account of truth values regarded as objects and, in part using this account, a solution to the problem of truthmaking for negative truths. Finally, **Chapter 6**, *The definability of negation as impossibility*, investigates various languages capable of defining negation in its strongest sense, viz. as impossibility where the modality in question is global **S5** necessity.

Chapter 2

What is negation?

Abstract

Logical revisionism is the doctrine that logic is revisable. Quine's famous objection to the possibility of logical revision is discussed and a response to that objection is given. The version of the objection I consider is restricted to negation in particular and contends that the notion of logical negation is not revisable. My response to that objection provides an analysis of negation according to which there is genuine disagreement between classicist and deviant on the nature of logical negation.

2.1   Introduction

What is negation? Seuren claims that it is "[j]ust about the most central operator in any logical system" ([Seu10, p. 31]) and Heinemann that it is

> one of the most primitive elements of human thought. There is no language without symbols, like no, not, none, etc. There is no system of logic, mathematics, science, philosophy or theology in which negation does not play a fundamental role. Generally speaking, no system, i.e., a coherent series of propositions referring to a specific subject, is possible without it, because its omission would destroy this very coherence. [Hei43, p. 127]

We need not go as far as Heinemann to appreciate the importance of negation and its figuring prominently throughout the history of logic and philosophy. The importance of characterizing negation—and here I do not presuppose there is but one such operation or expression deserving of the title—was made especially prominent when Quine famously claimed "Change of logic, change of subject". A lot of the time, differences between logics can be viewed as differences between their (alleged) negations in which case Quine's motto reduces to "Change of negation, change of subject". Disagreement between classicist

and deviant would be impossible if (i) each party gives different semantic analyses of negation, (ii) analyses differing in certain respects must be analyses of different things, and (iii) dialetheic and classical analyses of negation differ in just these respects. This raises two questions:

1. Do the deviant and classicist give different analyses of negation?

2. If so, do those analyses differ in respects that would preclude disagreement?

Before I attempt an answer to these questions—(1) in the affirmative *and* negative (or if one prefers, neither) and (2) in the negative—there are two important things that must be said about analyses. The first is that an analysis may fail to *uniquely* determine whatever it is intended to be an analysis of (e.g. a meaning, concept, object, etc.). I'll call such analyses *non-uniquely determining*. Non-unique determination does not in itself make an analysis *defective*. A typical proof-theoretic analysis of negation in terms of natural deduction rules analyzes negation as any connective satisfying the following rules:

$$
\begin{array}{cc}
\begin{array}{c}
[A] \\
\vdots \\
\dfrac{\bot}{\neg A} \, (\neg\mathrm{I})
\end{array}
&
\dfrac{A \qquad \neg A}{\bot} \, (\neg\mathrm{E})
\end{array}
$$

Yet there is a clear sense according to which such an analysis determines, if it determines anything at all, more than one propositional, and not necessarily *truth*, function. The intended classical one obeys the usual recursive truth conditions, viz.

- $v(\neg A) = \mathfrak{t}$ iff $v(A) = \mathfrak{f}$

and others do not; in particular, there is a propositional function consistent with the above analysis such that for some valuation $v_{\mathfrak{t}}$, $v_{\mathfrak{t}}(A) = \mathfrak{t}$ for all A whence, in particular, $v_{\mathfrak{t}}(\bot) = v_{\mathfrak{t}}A = v_{\mathfrak{t}}\neg A$.[1] In other words, $(\neg\mathrm{E})$ and $(\neg\mathrm{I})$ are sound when $\neg$ is interpreted as the usual truth function $\neg_1$ such that $\neg_1(\mathfrak{t}) = \mathfrak{f}$ and $\neg_1(\mathfrak{f}) = \mathfrak{t}$ and also when it is interpreted as the constant truth function $\neg_2$ such that $\neg_2(x) = \mathfrak{t}$ for $x \in \{\mathfrak{t}, \mathfrak{f}\}$ *under the important assumption that $\bot$ be given a nonstandard interpretation*, namely one that allows it to take the value $\mathfrak{t}$ under some valuation. This phenomeon has been labelled the "categoricity problem".[2]

Should we criticize the above analysis as being deficient? Matters are subtle, and I shall not go into the details here, but unique determination is not typically held to be a necessary condition for the adequacy of analysis. Consider (first-order) Peano arithmetic, **PA**. That it fails to uniquely pin down its intended model does not make it inadequate. Rather (on the assumption that

---

[1] What does it mean for an object, in this case a propositional function, to be consistent with an analysis? An answer could be given in detail but for present purposes it suffices to give the bare essentials. Here all that is meant is that any consequence relation closed under $(\neg\mathrm{I})$ and $(\neg\mathrm{E})$ is characterized by (i.e. sound and complete with respect to) a semantics defined relative to a class of valuations including $v_{\mathfrak{t}}$.

[2] See e.g. [Smi96], [Rum97], [Rum00], and [Hum00].

categoricity is desirable) it is not so much a fault of the analysis as a fault of the language. For given the expressive power of a first-order language (in the sense of Lindström theorem), **PA** is the best we could hope for as an analysis of informal arithmetic (save for the desire to, e.g., capture soundness internally by the inclusion of reflection principles).

Moreover, non-categoricity need not pose a problem for inferentialism since the meaning of a connective will not be given by a model-theoretic semantics. The fact that the proof theory does not link up with the model theory will not be a worry for an inferentialist who denies that any such link is desirable. What would be troubling is if a given proof-theoretic characterization of an n-place expression $\otimes$ failed to uniquely determine, relative to a formal theory **L**, that expression in the sense that $\otimes$ and $\otimes'$ do not figure in the same deductions, e.g. if $\otimes(A_1, \ldots, A_n) \vdash_{\mathbf{L}} B$ but $\otimes'(A_1, \ldots, A_n) \nvdash_{\mathbf{L}} B$. This is not the case for negation characterized according to $(\neg E)$ and $(\neg I)$ since we have, for any two connectives $\neg$ and $\sim$ satisfying, respectively, $(\neg E)$ and $(\neg I)$ and $(\sim E)$ and $(\sim I)$, we have $\neg A \vdash \sim A$ and conversely. In this sense the introduction and elimination rules for negation uniquely characterize a unary operator taken in a *proof-theoretic* sense. We might take this as justification of the proof-theoretic analog of the Fregean thesis that sense uniquely determines reference.

I conclude that non-unique determination of analysis does not preclude adequacy of analysis. This brings us to a second important fact about analyses: not only are they non-uniquely determining, but distinct analyses may be analyses of one and the same thing. Myriad examples abound: Turing, Gödel and Church's analyses of computability provide distinct analyses of the class of general recursive functions.

The chapter proceeds as follows. In section 2.2 I discuss the plausibility of a syntactic characterization of negation and conclude that such analyses are doomed to failure. In the following section 2.3 I go on to consider semantic characterizations, their difficulties and their advantages over a syntactic characterization.

2.2 Syntactic characterizations of negation

Let us think of negation as an expression of a given (possibly uninterpreted) language. We may then ask what constitutes an expression's being a negation. One way to settle the question is to lay out (relative to some background logic) inferential rules (under which axioms are subsumed as zero-premise rules), admissible or derivable,[3] which *characterize* the inferential behavior of the expression and which, moreover, are intended to be *characteristic of* a negation. If one is an inferentialist, then these rules may be said to bestow meaning, either in part or in full, on the expression and the expression may be said to have as its extension or intension that of a negation. Otherwise the rules may (but need not) be thought to be those sound relative to some intended semantics for the negation expression in question.

Nothing is assumed about the rules, e.g. that they be recursive. It is typically assumed that rules be finite and recursive, the latter in the sense that,

---

[3]A rule is admissible if whenever each of its premises is provable from no hypotheses, then so is the conclusion, and it is derivable if the conclusion is obtainable from the premises by finitely many applications of the primitive rules.

when the rule is taken as a set of n-tuples each of which is recursively associated with a natural number, the set of corresponding naturals is the range of a recursive function. One might find it useful, however, to forgo the recursiveness assumption for a number of reasons. For example, consider rules defined over uncountable languages even when (well-formed) formulae are always finite. In such a case even intuitively computable rules such as 'from $A \wedge B$ infer A' would count as non-recursive since the function, taken set-theoretically, is uncountable and is thus not the range of any recursive function.

If we look at the history of logic one finds prominently each of the following rules (or axioms) as being thought (partly) characteristic (or some subset as being wholly characteristic) of a negation $\neg$.

$$\frac{\overset{[A]}{\underset{\vdots}{B}} \quad \overset{[A]}{\underset{\vdots}{\neg B}}}{\neg A} \text{ (IR)} \qquad\qquad \frac{\overset{[\neg A]}{\underset{\vdots}{B}} \quad \overset{[\neg A]}{\underset{\vdots}{\neg B}}}{A} \text{ (R)}$$

$$\frac{A \to B \qquad \neg B}{\neg A} \text{ (IContra)} \qquad\qquad \frac{\neg A \to \neg B \qquad B}{A} \text{ (Contra)}$$

$$\frac{\neg(A \wedge B)}{\neg A \vee \neg B} \text{ (DeM1)} \qquad\qquad \frac{\neg A \vee \neg B}{\neg(A \wedge B)} \text{ (DeM1)}$$

$$\frac{\neg(A \vee B)}{\neg A \wedge \neg B} \text{ (DeM3)} \qquad\qquad \frac{\neg A \wedge \neg B)}{\neg(A \vee B)} \text{ (DeM4)}$$

$$\frac{A \qquad \neg A}{B} \text{ (EFQ)} \qquad\qquad \frac{A \qquad \neg A}{\neg B} \text{ (MEFQ)}$$

$$\frac{}{\neg(A \wedge \neg A)} \text{ (LNC)} \qquad\qquad \frac{}{A \vee \neg A} \text{ (LEM)}$$

$$\frac{\neg\neg A}{A} \text{ (DNE)} \qquad\qquad \frac{A}{\neg\neg A} \text{ (DNI)}$$

$$\frac{}{\neg(A \to \neg A)} \text{ (Aristoteles)} \qquad\qquad \frac{A \to B}{\neg(A \to \neg B)} \text{ (Boethius)}$$

Here MEFQ stands for 'Minimal Ex Falso Quodlibet', IR for 'Intuitionistic Reductio', IContra for 'Intuitionistic Contraposition' and DeM for 'De Morgan'. The last two principles, Aristoteles and Boethius, are principles of connexive logics and are inconsistent with classical logic (when taken *schematically*[4]). The rest of the principles, on the other hand, are all classically warranted.

The list is not intended to be exhaustive but I think it is nearly so.[5] More-

---

[4]A, B, etc. are metavariables which range over sentences of an unspecified object language. As they occur throughout this essay they are to be given the generality interpretation in which case all of the aforementioned principles (EFQ, IR, etc.) are to be understood schematically.

[5]I have left out principles involving quantification and non-truth-functional operators such as modal ones. For instance, typically the duality of the modal operators $\Box$ and $\Diamond$ is assumed—i.e. it is assumed that $\Box A \equiv \neg\Diamond\neg A$ and $\Diamond A \equiv \neg\Box\neg A$—and this is due partly to the behavior of negation. There are just too many principles to consider in a sufficiently rich language, so instead I shall focus on a propositional language containing just the usual (not necessary truth-functional) connectives such as $\to$, $\vee$ and $\wedge$.

over some of the principles are derivable from others given certain assumptions regarding structural properties such as cut. (I will be assuming that the premises are always bunched by set-theoretic union and that 'A' is to be read '$\{A\}$'.) Some of the principles, e.g. Contra, are "interaction principles" in the sense that they describe the way negation interacts with other connectives such as implication. Interaction principles depend just as much on the general behavior of each of the interacting connectives that figure in them as they do on the general behavior of negation. This does not, however, make them any less principles *characteristic of* negation.

One may find logics with so-called "negations" (and I shall not always use quotes to distinguish alleged from genuine negations, for then I would have to *always* use quotes lest I beg the question against anyone) that fail to satisfy at least one of the principles from the list. For example, classical negation fails to satisfy both of the connexive principles; intuitionistic negation fails to satisfy DeM1, DNE, LEM, Contra, R; and some paraconsistent (e.g. dialetheic) negations fail to satisfy IR, IContra, MEFQ (and hence EFQ) and LNC. If we take this at face value, it seems that none of the conditions listed is necessary for an operation's being a negation. But then either we must look for necessary conditions formulated, not in terms of rules, but in terms of *failures* of rules (e.g. never should $A \vdash \neg A$ for all A), or we must forget necessary conditions altogether and opt for a partial characterization via sufficient conditions.

Lenzen [Len96, p. 40] takes the former route (among others) and gives four principles he deems as being "*[u]nacceptable* principles which a logic **L** must never satisfy, if its "negation" operator $\sim$ is to rate as a real negation":

$$\frac{A \wedge B}{\neg A} \text{ (UN 1)} \qquad\qquad \frac{A \vee B}{\neg A} \text{ (UN 2)}$$

$$\frac{A}{\neg A} \text{ (UN 3)} \qquad\qquad \frac{}{\neg A} \text{ (UN 4)}$$

If $\wedge$ and $\vee$ have their usual properties, then UN 1-3 are equivalent. Moreover if EFQ holds then each is equivalent to the others. Each appears unacceptable since if a logic **L** (satisfying certain minimal conditions such as cut and adjunction, i.e. if $\vdash_{\mathbf{L}} A$ and $\vdash_{\mathbf{L}} B$ then $\vdash_{\mathbf{L}} A \wedge B$) has any theorems, then by UN i ($i \leq 4$), **L** is negation inconsistent. For let A be a theorem. Then by adjunction, $A \wedge A$ is a theorem, so by UN 1 (and cut), $\sim A$ is a theorem. Dually one derives the same consequence using $\vee$ introduction and UN 2.

There are two problems with requiring that a negation be *characterized* as not satisfying any of the UN i. The first is that they appear not to be *characteristic* of negation even if any negation operation must satisfy them. Similarly, we should not think that $x + x = x + x$ is characteristic of addition even if it is a principle that any addition operation must satisfy. Now if the hope is to amass enough necessary conditions to fully characterize negation, I do not see that the UN i should be included among those conditions.

The second, and more pressing problem, is that there are expressions arguably worthy of negative status that need not fail to satisfy some or all of the UN i. For consider a negation $\sim A := \Diamond \neg A$, considered by Lenzen himself under the guise of "weak negation", which applies only to contingent sentences A.[6]

---

[6]See also [Bèz02] in which the same negation, in the context of **S5**, is taken to be paraconsistent and so **S5** is argued to be, at least implicitly, a paraconsistent logic. The logic

Then for any contingent sentences A, B, it holds that $A \wedge B \vdash \sim A$, $A \vee B \vdash \sim A$ and $A \vdash \sim A$. In other words, satisfying either UN 1, UN 2, or UN 3 do not appear to be necessary for being a negation, at least if we have not applied sufficient constraints on the interpretation of our language.

Now if the UN i are to be taken unrestrictedly such that the metavariables range over arbitrary sentences or propositions rather than e.g. contingent ones, then in the example just considered one may take as axiom $A \to (\Diamond A \wedge \Diamond \neg A)$ to ensure that each formula, including tautologies and contradictions, is provably "contingent" if true. A more radical view, taken up by certain conventionalists, is to hold that $\Diamond A$ for each A. For instance Mortensen [Mor89] advocates the thesis of possibilism, by which he means "the group of theses that all propositions are possible, or possibly true, that all propositions are contingent, that no proposition is necessary". For such conventionalists a definition of negation as $\sim A := \Diamond \neg A$ will come with an endorsement of UN 1-3. Thus there are theories (e.g. some forms of conventionalism) according to which certain UN i hold unrestrictedly and so they ought not to be taken as always failing.

Here are two more examples that pose considerable strain on the idea that negation can be given a purely syntactic characterization. Consider a dialetheist such as Priest who argues (e.g. in [Pri99]) that the (purported) negation $\neg_{\mathbf{LP}}$ of $\mathbf{LP}$ (from hereon I shall denote *the* purported negation of a logic $\mathbf{L}$ by $\neg_{\mathbf{L}}$ when the logic contains exactly one negation) is a contradictory-forming operation on sentences or propositions. Clearly, given his semantics, it is not a contrary-forming operation, so how then can it be a contradictory-forming (i.e. contrary- and subcontrary-forming) one? His claim rests on the fact that $\neg_{\mathbf{LP}}$ satisfies both LNC and LEM. But is that enough? Let $\neg A$ be read 'It is such that $1 = 1$ or A—i.e. $\neg$ is a "trivial operator" in the sense that $\neg A$ for all A. Then $\neg$ satisfies most of the negation properties from our list above, including LNC and LEM. But, I claim, there is no plausible sense according to which $\neg$ so defined is a contradictory-forming operation and this speaks against Priest's claim that the mere satisfaction of LNC and LEM is constitutive of an operation's being contradictory-forming.[7]

Finally consider the "negation" of minimal logic defined by $\neg A := A \to \phi$ for a given ϕ *assigned no special property*. In effect, ϕ behaves just like an arbitrary atom. The mere fact that minimal "negation" satisfies certain "negation rules" such as IR (Intuitionistic Reductio) does not *by itself* make it a negation. There is just no reason, for instance, to think that 'If snow is white then Socrates is male' is a, let alone *the*, negation of 'Snow is white'. While implication to an "unwanted" formula may constitute the negation of

---

of the negation defined by $\sim A := \Diamond \neg A$ for $\mathbf{S5}$-$\Diamond$ is not closed under provable implications (and hence equivalents). Compare this with the empirical negation of chapter 3 which may be thought of in a similar light as being defined in a hybridized $\mathbf{S5}$ as $s(@, \neg A)$ where s is a satisfaction operator having the following truth conditions: $M, a \models s(b, A)$ iff $M, b \models A$. It too fails to be closed under provable implications.

[7] Priest does maintain (in personal communication) that LNC and LEM, even if not sufficient, are at least necessary conditions on being contradictory-forming but I have yet to see any convincing argument for the claim. Certainly an intuitionist who holds that falsity is *refutability* (rather than unprovability which may be undecidable relative to a given formal system $\mathbf{L}$) would disagree, for on such a view both A and $\neg A$ cannot be refutable (i.e. false) and so $\neg$ would be subcontrary-forming even though it fails to satisfy $A \vee \neg A$. Consider also Aristotle himself who likely rejected the universal validity of LEM on account of future contingents but who did not, as far as I can see, reject propositional negation as being contradictory-forming and predicate negation as being at least contrary-forming.

that formula depending on what is meant by 'unwanted', there is no reason to think that ϕ, in the context of minimal logic, is unwanted in the required sense.

Gabbay and Hunter [GH99] define a negation as implication to a formula from a given set of "unwanted" formulae. Typically, as in the case of intuitionistic and classical logic, the set is just $\{\bot\}$ where $\bot$ is a formula satisfying EFQ. In minimal logic, the same sign $\bot$ is typically used but it is not assigned any special properties which set it apart from an arbitrary atom. There is thus no reason to take it as an unwanted formula, and hence no reason to take $\neg A := A \to \bot$ as defining anything in the vicinity of negation in the sense of [GH99]. Like considerations apply to weaker negations such as the subminimal negation of [Haz95].

The lesson to be drawn from these three examples is that looking to characterize negation purely by syntactic means seems a hopeless enterprise unless we constrain the (possibly informal) interpretations of our language. Semantic considerations will therefore need to play a role in characterizing negation even when that characterization is mainly proof-theoretic. (Other considerations may play a role too, such as adequacy to linguistic data or pragmatics.) Note that a proof-theoretic characterization of negation may be semantic in the required sense. By 'syntactic' I mean 'completely devoid of semantic interpretation', and if one is a proof-theoretic semanticist, 'proof-theoretic' and 'syntactic' will then be far from synonymous.

2.3 Semantic characterizations of negation

The traditional conception

The traditional (or "toggle") conception ('TC' for short) of negation has it that

(TC) $\neg A$ is true iff A is false.[8]

On the usual assumptions about truth and falsity (and the metalinguistic biconditional), TC provides a classical account of negation, and by dropping any of these usual assumptions we end up with different accounts. Among the usual assumptions are the following two semantic properties:

(Bivalence) Each sentence is either true or false;

(Contravalence) No sentence is both true and false.

Dropping contravalence gives us a glutty theory of negation and dropping bivalence gives us a gappy theory.

Yet there is a lot more to the semantic characterization of negation than just these three properties. For instance, if we permit more truth values than just truth and falsity then we can distinguish two operators satisfying TC in a language satisfying both bivalence and contravalence. So these three properties do not in themselves *uniquely* characterize a single operator that we might wish to call 'negation'. For suppose we have a language L satisfying bivalence and contravalence and we have three truth values, truth ($\mathfrak{t}$), falsity ($\mathfrak{f}$) and

---

[8]This "toggle" view of negation is endorsed e.g. by Priest [Pri93] and Smiley [Smi93]. It is the most widely endorsed definition of negation.

indeterminacy ($\mathfrak{i}$). Then, as nothing in TC, bi- and contra-valence prohibits negation from being *non-functional*, we can define two distinct negations satisfying TC, one, letting $r \subseteq Form \times \{\mathfrak{t}, \mathfrak{f}, \mathfrak{i}\}$, which prohibits $r(A, \mathfrak{t})$ and $r(A, \mathfrak{i})$ from holding together and the other which prohibits $r(A, \mathfrak{f})$ and $r(A, \mathfrak{i})$ from holding together.

So if TC characterizes negation, it can only do so in part even in the presence of relatively strong restrictions (viz. bi- and contra-valence) on the language. This is an interesting fact given that some authors assume that these three conditions, viz. TC, bi- and contravalence, characterize negation uniquely. For instance, Copeland thinks they characterize classical negation:

> What the classical logician means by negation is exhausted by the statements that (1) every sentence A dealt with is either true or false (but not both), and (2) the negation of A is a sentence which is true whenever A is false, and false whenever A is true. Thus an applied semantics assigns the meaning of classical negation to a symbol $\neg$ just in case the semantics distributes truth values in accordance with (1), and embodies the condition that whenever A is true, $\neg A$ is false, and vice versa. [Cop86, p. 485]

Now one might have thought that Copeland was working under the background assumption that only truth and falsity are genuine logical truth values, in which case we do not yet have a counterexample to Copeland's claim. But even under this assumption his claim does not hold true unless truth and falsity are *unique*. While this is typically assumed it is by no means uncontentious. One may think of designationhood and undesignationhood as truth and falsity, respectively. Indeed Suszko held something of this view. But then there is no reason to think that either truth or falsity is unique. We may think, instead, that each designated value has just as much the right to being called 'truth' as any other, and indeed that each *is* truth, or that each undesignated value has just as much the right to being called 'falsity' as any other. This view is unproblematic if truth and falsity are properties rather than objects, but if they are objects in their own right, then two values can be truth or falsity only if 'be' is picking out some kind of many-one identity relation. I do not think, however, that many-one identities have any role to play in individuating truth and falsity.

In any case, we still run into problems even if truth and falsity are unique. For whether there is a unique operation picked out satisfying Bivalence, Contravalence and TC depends on what properties 'truth' and 'falsity' have and whether the metalanguage in which the three principles are stated is classical or not. A dialetheist who uses a dialetheic metatheory may hold that her semantics is bi- and contravalent and that her negation satisfies TC even though that negation is not recognizably classical. Or consider a dialetheist who, in some sense, dismisses the distinction between object- and metatheory (e.g. Priest). (One might view this equivalently as simply maintaining the distinction while opting to use the same theory for both.) Then if she expresses 'A is false' by '$\neg A$ is true' and her accepted logic has as theorems LEM and LNC (as e.g. **LP** does) then, at least by her own lights, she endorses bi- and contravalence.[9]

---

[9]In this case the classical three-valued functional semantics (rather than the two-valued relational semantics) for **LP** would have to be discounted by the dialetheist as intended or

However the issue is whether she can convince those who do not endorse the same metatheory that her semantics really is bi- and contravalent. Offering up a semantics couched in a classical metatheory which refers to entities to be regarded as truth and falsity and which includes interpretations that assign to sentences *both* truth and falsity is not going to do any such convincing. While it may help a classicist make some sort of sense of the dialetheist's reasoning, it will not convince the classicist that 'not' in the mouth of the dialetheist means *not*, or that her semantics really is contravalent. The only way to do this would be to convert the classicist over to dialetheism or to convince her that another account of contravalence is to be preferred, one on which dialetheic negation turns out to be genuinely contradictory-forming.

In what follows I argue for an account of contrariety and subcontrariety, and hence contradictoriness, on which a number of non-classical negations turn out "legitimate" when negation is regarded as a contradictory-forming operation. Thus if there is truth in the idea that logicians of various stripes share the core meaning of 'not' and are hence not speaking past each other, that idea is supported by the following account of contradictoriness. Conversely, the idea of shared meaning goes toward supporting the adequacy of the present account, so support feeds in both directions.

Contrariety and subcontrariety

It is often assumed that there is only one type of negation—the contradictory-forming type. Indeed it is thought that there is not just one *type* of negation, but just one negation altogether. The alleged uniqueness of negation comes from the uniqueness of contradictories and justifies the apparent correctness of locutions of the form "*the* negation of A". The same is not true of contraries and subcontraries which are not unique, in which case talk of *the* contrary or subcontrary of a given proposition is infelicitous.[10] For these reasons, and others to be discussed below, if negation is a unique operation on sentences (or propositions) then it cannot be a contrary- or subcontrary-forming operation. The only plausible candidate left, then, is that negation is a contradictory-forming operation. But since contradictories are unique, it must be *the* contradictory-forming operation.

I wish to take these ideas seriously, namely, that contradictory-forming operations are negations and that contradictories are unique. There are many reasons to do so. First, I have never encountered anyone who rejects the uniqueness of contradictories. Second, the view that contradictory-forming operations are negations is widely accepted and hence uncontroversial. Moreover on a compelling view of contradictoriness (viz. the "prime account" below) a whole host of "negations" turn out genuine. This is good, as there are at least *prima facie* grounds to think that the deviant and classicist are really disagreeing about the nature of negation when, for example, the classicist endorses LEM and the intuitionist refrains from any such endorsement.

---

faithful. They might hold instead that it serves only to convince the classicist that their view is coherent in a minimal sense. In any case the functional semantics, arguably, leads the dialetheist into accepting trivialism (see [Smi93] for such an argument), the thesis that everything is true, so they should prefer the relational semantics anyway.

[10]Geach [Gea69] criticizes McCall [McC67] for making just this "uniqueness" mistake.

Every A is B ... No A is B

A —— contraries —— **E**

subaltern ... contradictories ... subaltern

I —— subcontraries —— **O**

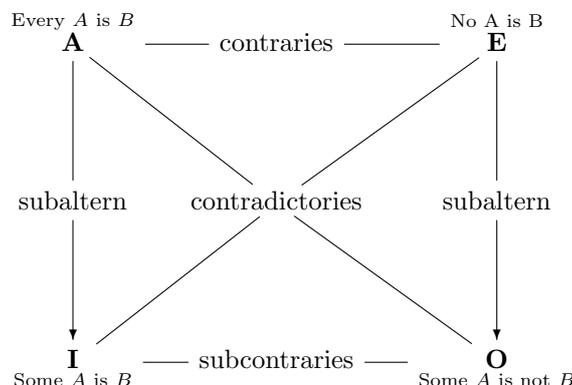Some A is B ... Some A is not B

Figure 2.1: Square of opposition

The Aristotelian and one-sided accounts

According to the *modern account*,

- two sentences are contraries (of each other) if both of them cannot be true together;

- two sentences are subcontraries if both of them cannot be false together.

This has not always been the received view. The following account was propounded by Aristotle and is not equivalent to the modern account on several interpretations:

- the A (positive universal) and E (negative universal) form sentences are contraries;

- the I (positive existential) and O (negative existential) form sentences are subcontraries.

An A form sentence is one of the form 'Every A is B', an E form sentence is one of the form 'No A is B', an I form sentence is one of the form 'Some A is B' and an O form sentence is one of the form 'Some A is not B'. They constitute the four corners of the Aristotelian "square of opposition" (see Figure 2.1).

The "canonical" translation of Aristotle's A and E sentences, according to their *surface structure*, are $\forall x(A \to B)$ and $\forall x(A \to \neg B)$ respectively, while those of the I and O sentences are $\exists x(A \land B)$ and $\exists x(A \land \neg B)$ respectively. (Here I assume A and B contain precisely x free.) However, in classical model theory, A and E sentences may be true together (under a model M) when e.g. A has an empty extension (in M) and I and O form sentences may be false together when, again, A has an empty extension. We are assuming a translation according to surface structure and so it does not build into it the existential presupposition of the subject term A nor of the term B. This is, of course, controversial and building existential presupposition in does indeed circumvent the counterexample just given showing that A and E form propositions are simultaneously satisfiable when the subject term is empty. Circumvention is

also possible when, instead of building in existential presupposition, the notions of contrariety and subcontrariety are always given relative to the background assumption that the subject term (of the A, E, I and O sentences, being the same) has a non-empty extension.

Moreover, on the modern account two sentences may be both contraries and subcontraries, in which case they are contradictories, in contrast to a widely held account, propounded in numerous textbooks, which holds that

- two sentences are contraries if they cannot be true together but *they can be false together*;

- two sentences are subcontraries if they cannot be false together but *they can be true together.*

I shall call this account (not to be confused with the Aristotelian or modern account) *one-sided.* On this account, contradictoriness cannot be defined as the intersection of contrariety and subcontrariety and this seems to me to be a serious drawback since the relation between the contrariety, subcontrariety and contradictoriness becomes unclear or unnecessarily complicated (on which more later).

There are further complications besides. If contraries must be capable of being false together, then nothing is the contrary of a logical truth since a logical truth cannot be false. Similarly, nothing could be the subcontrary of logical contradiction. This was noticed by Sanford [San68] as a problem for the Aristotelian account according to which A and E sentences are always contraries and I and O sentences always subcontraries. But if some A sentences are logically true (such as 'Every green egg is green') then no sentence, including its corresponding E one, is its contrary and it follows that a one-sided account is inconsistent.

One possible emendation of the one-sided account is to hold that there is a presupposition that the A, E, I, O form sentences figuring in the account be *logically contingent* in which case the proposed counterexamples do not work. The *emended one-sided account* is then

- two sentences are contraries if (i) *each is logically contingent* and (ii) they cannot be true together but they can be false together;

- two sentences are subcontraries if (i) *each is logically contingent* and (ii) they cannot be false together but they can be true together.

But even such an emendation is susceptible to counterexample. An example due to Humberstone [Hum03] runs as follows. Let the relation of accompaniment between x and y hold precisely when x and y are mereologically disjoint (concrete) worldmates (in the sense of Lewis [Lew86]). Then the A form sentence 'Every table is accompanied' can be neither true *nor false* together with its corresponding E form sentence 'No table is accompanied' even though each is contingent. For the falsity of the A sentence implies that some table is unaccompanied and the falsity of the E sentence that some table is accompanied, whence the two cannot be false together.

Humberstone's alleged counterexample is not unproblematic. For all it shows is that there are A and E form sentences that, *under some given interpretation*, cannot both be false. But if we are to read "can both be false" in

the emended one-sided account as meaning "there exists an interpretation to the non-logical vocabulary of the A and E form sentences under which they are both false" then Humberstone's example does not provide a *counter*example. For certainly we may interpret 'accompanied' and 'table' so that 'Every table is accompanied' and 'No table is accompanied' both turn out false: just interpret 'table' to mean table and 'accompanied' to mean green. But is this how we ought to read "can" in the emended account?

There are three plausible readings of 'can' in the emended one-sided account, the first of which was already mentioned:

1. there exists an interpretation to the non-logical vocabulary of the A and E form sentences under which they are both true;

2. taking the A and E sentences schematically, there are instances A and B of them that are true under some interpretation;

3. there is a *representation*, i.e. possible world, in which the A and E form sentences are true under a fixed interpretation of the vocabulary (such that 'table' refers to tables and 'accompanied' to the accompaniment relation) though the domain of discourse may vary, i.e. extension, but not intension, may vary;

The first and second come to the same thing and seem to me to be the most plausible reading of 'can' in the emended account since they follow from the assumption that each corner of the square is to be read schematically.

The issue of schematicity relates to a similar question regarding Aristotelian logic: Are the syllogistic rules *schematic* or not? If the widespread schematic interpretation of the syllogistic is correct then we might think the same holds true when interpreting the square of opposition, and this is reason to reject Humberstone's alleged counterexample to the foregoing emended account since it reads 'can' according to (3), i.e. non-schematically, rather than (1) or (2).

So we see that the emended one-sided account (whereby contraries must be contingent) works when the A and E sentences are taken schematically but not when they are taken as interpreted sentences. Yet one might think it *prima facie* compelling that an account of contrariety ought to apply also to interpreted sentences, suggesting that a satisfactory account of contrariety and subcontrariety is one that gets things right *however we interpret the A and E sentences*, i.e. whether schematically or not.

But if the language under consideration is interpreted then the modal 'cannot' which occurs in the accounts considered here has the force of the non-modal 'is not', for there would only be one interpretation to consider, viz. the one under which the sentences have been interpreted.

A notion of contrariety and subcontrariety which applies to fully interpreted sentences (i.e. ones which have their extension fixed) does not appear to be of much interest then for what it amounts to is the following:

- two sentences are contraries if they are not true together (and possibly some other conditions);

- two sentences are subcontraries if they are not false together (and possibly some other conditions).

But nobody should think that 'Every flower is red' and 'Every tire is round' are contraries simply because they are not true together (on account of the former).

Aristotle, but not all of his followers, have a way out of this worry if the worry is even worth escaping. He understands contrariety as a relation, not just between sentences with which we have been here concerned, but between predicates and properties (and some would think individuals and individual terms) too. If whiteness and blackness (or 'is white' and 'is black') are contraries, then we can account for contrariety holding between 'a is white' and 'a is black' in virtue of it holding between whiteness and blackness (or the predicates). But moreover, we can account for contrariety failing to hold between 'Every flower is red' and 'Every tire is round' in virtue of its failing to hold between the property of being a red flower and that of being a round tire. The only problem with this account is that there are no established criteria which tell us whether a given pair of properties (or predicates) are contraries—not a problem for Aristotle, but certainly a problem for any rigorous account of contrariety.[11]

We could not, for example, say that two predicates ϕ and ψ are contraries when the intersection of the properties they express is empty, at least when properties are taken extensionally. For being renate and not being cordate would then be contraries, but it is only "by chance" that renates are cordate. We could take properties intensionally in the sense of Lewis [Lew86], where a property is a set of possible world-bound objects, but then we have to buy into the metaphysical baggage of Lewis's theory. There are other options, but whichever one opts for, things do not appear to go as smoothly for contrariety taken as a relation between properties as they do for contrariety taken as a relation between sentences or propositions.[12]

Three accounts of contradictoriness

In providing a *general* account of contrariety and subcontrariety there is no reason to think logical truths and contradictions are *degenerate* cases to be excluded from the account. Indeed, there just is no reason to think that *any* type of sentence is degenerate as regards an account of these three properties which explains why the foregoing emended accounts are *ad hoc*.

---

[11]Indeed things may be more complicated if Aristotle held, in opposition to the modern account, that contraries be *unique* as argued in [Bog92]. For then Aristotle needs to give an account of why whiteness rather blueness is the contrary of blackness. Roughly the account he gives is that, on the color scale, the two are at polar ends. This might work for scales bounded on each end but it is not general enough to cover unbounded scales (e.g. those isomorphic to the open interval $(0, 1)$). The properties bigness and smallness are like this, as there is neither a maximal bigness nor maximal smallness and, in any case, one can always conjure up scalars denoting more extreme properties, such as 'extremely big', 'larger than extremely big', etc.

[12]Things go smoothly if we consider contraries formed by negating a predicate as in 'is red' and 'is non-red' but that leaves out a lot of an account of contrariety. However, it has been said in [Hor01, p. 205] that others have claimed that the negation of a scalar term does not yield a contradictory or contrary "opposition". This is clearly mistaken. The examples cited therein show only that for some pairs of sentences A and A', where A' is the result of negating a scalar term in A, both *may be false* or that, concerning the relevant terms, there is a "failure of the Law of Excluded Middle". But all that shows is that the negation of a scalar term does not always yield a *subcontrary*, and hence contradictory, opposition. It does not show it does not yield a contrary. Surely it does.

Let us, then, take as our default account of contrariety and subcontrariety the modern account discussed above, viz. that two sentences are

(SubCon) subcontraries if they cannot be false together;

(Con) contraries if they cannot both be true together.

McCall [McC67] adopted these definitions one year before Sanford [San68] endorsed the modern account as consistent and condemned the others as inconsistent, but they appear also long before both McCall and Sanford in Strawson's *Introduction to Logical Theory* (1952) and in Quine's *Methods of Logic* (1950).

Priest (in personal communication) gives the following alternative account. Two sentences are

(SubCon*) subcontraries if one is true whenever the other is not true;

(Con*) contraries if one is false whenever the other is not false.

Intuitively SubCon* says that $\sim A$ is the subcontrary of A in the sense that both cannot fail to be true and Con* says that $\sim A$ is the contrary of A in the sense that both can't fail to be false.

The negation $\neg_{\mathbf{LP}}$ (by which I mean the truth function of the three-valued semantics of [Pri79]) satisfies SubCon* and Con* but it fails to satisfy SubCon. Notice that the negations of **FDE** and (weak and strong) Kleene logic **K3** fail both SubCon* and Con* (and so does the exclusion negation of **K3**).[13] This might spell bad news for Priest's account of contrariety and subcontrariety, for it precludes from being a negation, among others, any function that is a fixed point for a *neither true nor false* value. But notice that $\neg_{\mathbf{FDE}}$ also fails both SubCon and Con each of which is satisfied by Kleene negation. So the original account (given by SubCon and Con) classifies $\neg_{\mathbf{K3}}$ but not $\neg_{\mathbf{LP}}$ as a contradictory-forming operation, Priest's account classifies $\neg_{\mathbf{LP}}$ but not $\neg_{\mathbf{K3}}$ as a contradictory-forming operation, and both accounts classify $\neg_{\mathbf{FDE}}$ as neither a contrary- nor subcontrary-forming operation.

There is a more encompassing account of subcontrariety and contrariety that meshes well with traditional accounts. Let us say that two sentences are

(SubCon′) contraries if one is true whenever the other is false;

(Con′) subcontraries if one is false whenever the other is true.

On this account $\neg_{\mathbf{FDE}}$ is a contradictory′-forming operation, but so too are $\neg_{\mathbf{K3}}$ and $\neg_{\mathbf{LP}}$. If there is appeal in each of these being contradictory-forming operations, for some pretheoretic notion of contradictoriness, then SubCon′ and Con′, which I'll call the *prime accounts* (of subcontrariety, contrariety and contradictoriness), appear to be adequate where the other accounts fail.

That $\neg_{\mathbf{FDE}}$ fails to be contradictory-forming on the original and star (Priest) accounts should be especially troubling for relevantists. For negation on the "Australian plan" (i.e. Routley semantics) satisfies none of the three accounts of contradictoriness just considered. However, when negation is conceived on

---

[13]I have in mind the usual four-valued semantics of **FDE** defined relative to the diamond algebra FOUR in which negation is a fixed point for the values n (*neither true nor false*) and b (*both true and false*).

the "American plan", according to which worlds are **FDE** matrices, nega-
tion turns out to be contradictory-forming *only* on the prime account. (Any
of the usual relevant logics that can be given a semantics on the American
plan can be given one on the Australian plan. See [Rou84] for details.) Any
relevantist who considers his negation, construed on the American plan, as
a genuine contradictory-forming operation should therefore have sympathies
with the prime account. Indeed, the account might serve as reason to prefer
the American plan!

There is more to be said about the advantages of the prime account. Prin-
cipally it classifies the "negations" of various logics as genuinely contradictory-
forming, and if contradictory-forming operations are necessarily negations, then
each these contradictory-forming operations turns out to be a genuine nega-
tion on the prime account. If this is correct then we have a solution to the
Quinean "speaking-past-each-other" problem since on the prime account the
classicist, dialetheist and gap theorist all express negation by 'not'. The same
considerations may be used to diffuse objections of the same ilk such as those
of [Sla95].

Are there compelling grounds, then, for holding the prime account over
others? One has already been mentioned, viz. that it diffuses the Quinean ob-
jection which makes disagreement between deviant and classicist impossible.[14]
The prime account classifies, among others, $\neg_{\mathbf{FDE}}$, $\neg_{\mathbf{LP}}$, $\neg_{\mathbf{CL}}$, $\neg_{\mathbf{K3}}$ and the
exclusion negation $\neg_e$ of **K3** and $\neg_{\mathbf{IPC}}$ as contradictory-forming operations.[15]
Those sympathetic with the view that there is fixed meaning attached to 'not'
in the mouth of deviant and classicist will find this fact an extremely compelling
reason to favor the prime account over others.

Another reason to favor the prime account is that it appeals only to truth
and falsity and, importantly, not 'not', a virtue lacked by both Priest's "star
account" and the original. If we wish to define negation as a contradictory-
forming operation then we are best to do so without invoking (in the meta-
language) the very concept of negation itself lest the account be charged with
vicious circularity. With this wish should come the assumption that, as it
appears in the prime account, falsity is not to be defined as truth of negation.

One might also wonder about Priest's motivations for favoring the star ac-
count over the original. It would seem the *principal* reason is that it classifies
his preferred logic's negation $\neg_{\mathbf{LP}}$ as a contradictory-forming operation, and
hence a negation on most accounts, without doing serious injustice to the orig-
inal account of contradictoriness. As such it diffuses the objection raised by
Slater in [Sla95] that $\neg_{\mathbf{LP}}$ is not contradictory-forming. In fact one equivalent
way of formulating the original account is given precisely by the converses of
SubCon* and Con* and the converses follow from the original account on the
assumptions of bi- and contra-valence and this might account for the reason
the two accounts feel strikingly similar. The same is true concerning the orig-
inal and the prime account. The moral is that if one is free to choose her mild
variation on the original account of contrariety and subcontrariety to fit her

---

[14]The objection was given up, in any case, by later Quine, being sympathetic to logical
revisionism, who held that "no statement is immune to revision" ([Qui51, p. 40]) and, in
particular, that "[r]evision even of the logical law of the excluded middle has been proposed
as a means of simplifying quantum mechanics" ([Qui51, p. 40]). But his objection remains
forceful even if one is not a radical revisionist.

[15]I assume here that, for the intuitionist, truth is provability and falsity is refutability.

tastes then that is all the more reason to think the prime account is at least as good as the others for it more often gives us a favorable verdict on the question of whether a given alleged negation is genuine.

What other reason might we have for favoring the original over the prime account? One might think that its use of conjunction and negation is less problematic than the prime account's use of implication. But why would she think that? Certainly truth-functional conjunction is no less problematic than truth-functional material implication and if we are allowed to assume that the original account invokes truth functional conjunction, rather than a "problematic" intensional conjunction, then we are likewise allowed to assume that the prime account invokes unproblematic material implication. Moreover, imagine a linguistic community that does not have an expression for picking out conjunction in their language. Surely we should not think they ought to prefer formulations of contrariety and subcontrariety in terms of conjunction or that they lack the linguistic resources for expressing such notions as contrariety and subcontrariety for these notions appear not to be inextricably linked to conjunction.

2.4 Final remarks

In order for two parties to disagree on some subject matter S, it is not required that they attach precisely the same meaning to every word concerning their talk about S. If they did, they could not be disagreeing. All that is required for disagreement is that there be *enough* shared content or meaning in their disagreement talk. What counts as "enough" is a difficult matter to settle and not one that needs to be settled here.

I have argued that the deviant and classicist may hold the same *definition* of negation as 'a contradictory-forming operation'. Now there are at least two things we may conclude from this:

1. the deviant and classicist mean the same thing by 'not';

2. they do not mean the same thing by 'not' but there is enough shared meaning between them to allow for disagreement.

The first conclusion strikes me as highly implausible. We cannot mean the same thing by a word just by holding that the same string of symbols is its definition. If I think 'man' refers to donkeys then we do not mean the same thing by 'bachelor' even if we both agree that the sentence 'Bachelors are unmarried men' serves as a definition of 'bachelor'. When I claim that bachelors are four-legged and you object, we are not disagreeing, we are merely speaking past each other.

If the lesson to be learned comes in the form of the second conclusion, I confess I have not gone far enough in this end. For I have not argued that there is enough shared content or meaning for disagreement in the deviant and classicist's respective uses of 'true' and 'false', the key terms used in defining negation as a contradictory-forming operation. I leave this for another occasion, but I would like to add two further important points.

First, consider the classicist and gap theorist, the latter being one who rejects the law of excluded middle in the form $\vdash A \lor \neg A$. They may disagree

about which inferences involving negation are valid without disagreeing on truth or falsity or even the definition of negation. Even the classicist may hold, like the gap theorist, that truth and falsity are not exhaustive, i.e. that bivalence fails. However, they may hold that a sentence A follows from some others Γ just in case A is never false when each of Γ is true. That is a perfectly reasonable notion of consequence but it is one that can yield classical logic *on a non-bivalent semantics* whereas the standard one, according to which A follows from Γ just in case A is true whenever each of Γ is, would yield a gappy logic such as **K3**.[16]

Second, one might think that the meaning of negation just is its referent (under a direct reference theory of the meaning of logical vocabulary) and that the classicist and deviant are able to disagree about negation—the propositional function "out there" in the world—simply by some external relation holding between them, the world and language. What they are disagreeing about, then, is which one of their respective theories is getting things right. I have not been assuming this sort of externalist picture of the meaning of logical vocabulary because it takes too much for granted and it is much more difficult to justify when the relevant vocabulary has abstract objects as referents whose natures are the very thing in dispute. It seems the only plausible response to the Quinean objection is to assume a non-externalist picture of the logical vocabulary and hope to establish that there is some shared meaning between deviant and classicist in the manner I have attempted here.

So there may turn out to be a number of differences on which a particular case of disagreement hinges. I think when it is located in notions as fundamental as truth and falsity rather than the definition of negation itself, it appears more likely that the deviant and classicist do share enough meaning to disagree about negation when they agree on the definition but not on the fundamental notions such as truth and falsity, and that is what I have hoped to show here.

---

[16]Take the class of strong Kleene matrices and define consequence the first way such that A follows from Γ just in case A is never false when each of Γ is true. Consequence is then classical even though the semantics is not bivalent. The technical point is simple but its philosophical importance runs much deeper.

Chapter 3

Constructivism and empirical negation

Abstract

Intuitionistic negation, typically defined as implication to absurdity, is too strong for expressing claims with empirical content such as 'Goldbach's conjecture is not decided at present'. For what is claimed here is not that there is a procedure for taking any (purported) proof of 'Goldbach's conjecture is decided at present' to a proof of an absurdity but rather the weaker claim that the conjecture has not yet been decided. Thus an extension of constructivism to empirical discourse, a project most seriously taken up by Dummett and Tennant, requires an empirical negation lying somewhere between classical negation ('It is unwarranted that...') and intuitionistic negation ('It is refutable that...'). I put forward one plausible candidate for empirical negation that has a close affinity to classical negation. The present proposal is compared favorably to some others that have been propounded in the literature.

3.1 Introduction

In mathematical discourse a uniform treatment of negated and unnegated statements can be given by defining the former in terms of the latter. If A is a mathematical statement then its negation may be defined as $A \to \bot$ where $\bot$ is either taken as primitive or as an abbreviation of some fixed absurdity such as $0 = 1$. What counts as an "absurdity" will depend on the background formal theory. Notice that in the usual arithmetical setting, defining the negation of A as $A \to 0 = 1$ gives us the constructive properties of negation we expect. For example, *ex falso quodlibet*—that from A and $\neg A$ anything follows—is derivable from modus ponens and the fact that $0 = 1$ implies everything. Negation introduction and elimination are then just special cases of implication introduction and elimination.

However, there are issues concerning the choice of absurdity. If a theory does not contain sufficient arithmetic then obviously $0 = 1$ will not do; i.e. there may be no single sentence able to play the role of absurdity across all mathematical discourses. Instead we shall have to choose, for a given mathematical theory, some sentence able to assume the role of absurdity. In practice this poses no problem on the assumption, and this assumption seems safe enough, that we can effectively choose a suitable absurdity for any given theory we are likely to care about. For example in arithmetic we may choose $0 = 1$, in the theory of strict linear orders $0 < 0$, and so on.

Constructively there is nothing problematic about implication interpreted according to the familiar BHK clauses nor of an absurdity such as $0 = 1$, so negation turns out unproblematic in mathematical discourse.[1] Is the same true for empirical discourse? Dummett thinks not, stating:

> Negation ... is highly problematic. In mathematics, given the meaning of if ... then, it is trivial to explain "Not A" as meaning "If A, then $0 = 1$"; by contrast, a satisfactory explanation of "not", as applied to empirical statements for which bivalence is not, in general, taken as holding, is very difficult to arrive at. Given that the sentential operators cannot be thought of as explained by means of the two-valued truth-tables, the possibility that the laws of classical logic will fail is evidently open: but it is far from evident that the correct logical laws will always be the intuitionistic ones. More generally, it is by no means easy to determine what should serve as the analogue, for empirical statements, of the notion of proof as it figures in intuitionist semantics for mathematical statements. [Dum96, p.473]

A blanket term for the analogue of proof for empirical statements is *warrant* or *verification*. One need not spell out a precise theory of warrant in formulating a semantics whose primary semantic values are warrants. Indeed there has been no precise spelling out of proof for the constructivist, since proof for them is taken as intuitive and not relative to a given formal theory, though this has not prevented the formulation of numerous semantics for constructive logics.[2] A similar point is made by Kleene regarding the realizability interpretation of intuitionistic number theory when he states "[t]he analysis which leads to this truth definition is not to be regarded as more than a partial analysis of the intuitionistic meaning of the statements, since it takes over without analysis, or leaves unanalyzed, the component of evidence" [Kle45, p. 110].

In what follows I shall be taking the notion of warrant or verification as primitive and assuming, moreover, that a naive extension of a constructively acceptable semantics to empirical discourse is one which replaces proofs as semantic values with warrants (perhaps with other necessary modifications

---

[1]It should be mentioned in passing that some (e.g. [CC00]) have objected to the use of $0 = 1$ in a definition of negation, though I shall not enter into this debate here.

[2]Taking proof as proof-in-**L**, for some formal theory **L**, has been thought to be problematic for constructivism for reasons having to do with Gödel's incompleteness theorems which seem to show that there would be verification- or proof-transcendent truths. See e.g. Martin-Löf's [Mar84, p. 11].

made as well).[34] The verification of a statement A might be taken to be the holding *in principle* of sufficient empirical evidence in support of A, where what counts as evidence is either contextually determined or domain-specific. For instance, seemings to John might count as evidence for 'John is hungry' whereas they may not, e.g. because John has a fever, for 'The temperature is above twenty degrees celcius'. In other words, a number of factors may serve to determine what counts as evidence for a given class of propositions. In the case of mathematics, proof then turns out to be a species of warrant.[5]

It is easy to see why the usual notion of constructive negation does not by itself suffice for expressing negation in empirical discourse. Suppose we wish to express that Goldbach's conjecture is undecided at present. According to the arrow-falsum definition, this statement is equivalent to 'If Goldbach's conjecture is decided (i.e. proved or refuted) at present, then $0 = 1$'. But this is far too strong: it states that it is *refutable* that Goldbach's conjecture is decided when all that is meant is that the conjecture *has not yet been* decided. To express such claims the constructivist needs a *weak* negation that, when appended to a statement, expresses that the statement *lacks warrant at present*. Such a negation has been referred to as 'empirical' in [Wil94]) and [DS06] and as 'factual' antecedently in [Hey71, p. 18].

Dummett proposes that for empirical discourse we treat verification and falsification on a par by taking them as *sui generis* notions. However, if we do so then we must give up a *uniform* treatment of the conditions under which a sentence is verified or falsified. On this, Dummett remarks:

> [w]e might regard the meanings of negations of numerical equations as being given directly in terms of the computation procedures by which those equations are verified or falsified: 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

---

[3]Williamson [Wil94] makes this suggestion regarding the BHK clauses in order to show that a semantics so extended cannot make sense of empirical statements of the form 'A may be undecided'. I briefly discuss his argument in section 3.7.

[4]While I have taken the notion of warrant as primitive, so that one may fill in their favorite theory of warrant in the discussion to follow, one should keep in mind that the following discussion reveals what basic properties warrant must have, i.e. not just anything goes for warrants. For instance, in a constructivist setting the notion of warrant is regarded as *monotonic* in the sense that if A is warranted by a particular state a and b stands in the relevant "inclusion" relation to a then A is warranted also at b. This does not entail that warrant is defeasible. That would only be justified on an implausible reading of the inclusion relation between states.

[5]What counts as constructively acceptable is unclear. For instance it is not clear that disjunction ought to satisfy the disjunction property, viz. that if $A \vee B$ is warranted (recalling that the A and B may be empirical) then either A is or B is. Indeed one might question the tenability of the whole project of extending constructivism beyond mathematics. Even if the project is not tenable across the board, there is an interesting fragment of empirical discourse for which the project does appear tenable, viz. that which includes only mathematical statements and statements of the form 'There is no proof (or there is proof) of A at present' where A is mathematical. It is worth investigating the inferential role negation has even in this and similarly restricted settings.

> 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. [Dum96, pp. 71-72]

A semantics which treats verification and falsification symmetrically along the lines just sketched by Dummett has been proposed by Thomason [Tho69] for the strong negation of [Nel49]. Gurevich [Gur77] motivates his semantics for strong negation by observing that "[i]n many cases the falsehood of a simple scientific sentence can be ascertained as directly (or indirectly) as its truth. An example: a litmus-paper is used to verify sentence [sic] "The solution is acid" " ([Gur77, p. 49]) by which he means 'verify the falsity of' when he says 'verify'. In the context of constructivism we see that symmetrical treatments of truth and falsity have been around for quite some time.

Williamson [Wil94] suggests, but ultimately rejects, using strong negation as an empirical negation. I will not discuss the details of the argument here, but concerning his unargued claim that "anything recognizable as a negation operator in intuitionistic logic...will satisfy $[\sim A \rightarrow \neg A$ (where $\sim$ is any intuitionistic negation and $\neg$ the usual intuitionistic one)]" (p. 141), which he believes to be problematic for extending intuitionistic semantics to empirical discourse, I remark only that it flies in the face of the obvious fact that no negation at least as strong as intuitionistic negation is suitable for expressing claims (considered by Williamson in [Wil94]) such as 'Goldbach's conjecture is undecided at present'. This alone provides overwhelming reason to reject Williamson's claim concerning the strength of empirical negation.[6]

How ought an empirical negation behave? If we take the following quote of Dummett seriously then, at least relative the class of statements we have in mind, empirical negation ought to look classical or very nearly so:

> Our reluctance to say that pi was not transcendental before 1882, or, more signicantly, to construe mathematical statements are signicantly tensed, is not merely a lingering effect of platonistic misconceptions; it is, rather, to speak in this way would be to admit into mathematical statements a non-intuitionistic form of negation, as will be apparent if one attempts to assign a truth-value to 'pi is not algebraic', considered as a statement made in 1881. This is not because the 'not' which occurs in '...is not true' or '...was not true' is non-constructive: *we may reasonably view it as decidable whether or not a statement has been proved at a given time.* But

---

[6]The foregoing discussion applies equally well to an alleged solution of Wansing [Wan02] for evading Fitch's paradox since we have that the strong negation of a statement constructively implies the usual intuitionistic negation of that statement. Thus if intuitionistic negation is too strong to express the empirical negation of KA (where K is a knowledge operator used in expressing Fitch's paradox) then *a fortiori* so too is strong negation.

> though constructive, this is an empirical type of negation that oc-
> curs in statements of intuitionistic mathematics. [Dum77, p. 337,
> my emphasis]

If it is decidable whether or not a given statement A lacks warrant at present
(or any given time more generally), it will always be true, for example, that
'either A is (now) warranted or it is not'.

Empirical negation cannot simply be classical negation as there is no *straight-
forward* way of introducing classical negation to intuitionistic logic without
having the two negations collapse into classical negation. So on pain of col-
lapse, empirical negation must forgo certain classical principles. The same issue
arises within the setting of classical relevant logic (see [MR73] and [MR74]),
where the familiar law of contraposition in "arrow-form" fails (though it holds
in rule form), again, on pain of collapse. Precisely which principles our account
of empirical negation forgoes is discussed in section 3.4.

However, besides collapse, there is a host of other objections against clas-
sical and classical-like negations that any proposal for a constructively or rel-
evantly acceptable negation will have to avoid. In the next section I look at
these objections in turn and show that none of them are a problem for the
empirical negation to be introduced in section 3.3.

3.2 Constructivist objections to classical (and like) negations

Conservativity

If **L** and **L**$'$ are logics in the respective languages $\mathcal{L}$ and $\mathcal{L}'$ with $\mathcal{L} \subseteq \mathcal{L}'$, then
L$'$ is a *conservative extension* of **L** just in case for every sentence A in $\mathcal{L}$, if A is
L$'$-provable then it was already **L**-provable. Conservativity has been thought
to be a necessary condition on meaning coherence in the sense that the rules
governing a connective can only confer coherent meaning on that connective
if they conservatively extend a given coherent base.[7] There are a number of
reasons for desiring conservativity having to do primarily with anti-holism,
learnability, anti-realism and consistency (against e.g. tonk-like connectives),
but whichever reasons one has in mind, classical negation is going to be prob-
lematic since the usual ways of proof-theoretically extending deductive systems
for intuitionistic logic to include classical negation yield nonconservative ex-
tensions. A famous example witnessing this nonconservativity phenomenon is
Peirce's law, $((A \rightarrow B) \rightarrow A) \rightarrow A$, which is classically but not intuitionisti-
cally provable. As such, classical negation has been deemed incoherent, most
famously by Dummett.

We need not enter the debate about whether conservativity is a necessary
condition for coherence, since one can remain neutral on the issue in cases

---

[7]Dummett is often thought to have held the conservativity constraint. This view is criti-
cized by Read [Rea00] where Dummett is accused of confusing conservativity with harmony,
of which only the latter he seems to have explicitly endorsed. There are a number of criti-
cisms of the conservativity constraint which seem to settle the matter against it. One is that
there may be two connectives each of which can be individually and conservatively added to
a given logic though the addition of both yields a nonconservative extension. Is each of these
connectives individually coherent until present together? That seems an implausible thing
to say.

where a given connective meets the conservativity constraint in the first place. The empirical negation introduced in section 3.3 yields a conservative extension to intuitionistic logic, and so conservativity poses no problem for it.

Harmony

If conservativity fails as a requirement for coherence, the inferentialist will need to appeal to some other requirement to rule out classical negation as incoherent. Harmony has been thought to fill this role. Harmony is the idea that introduction and elimination rules must be harmonious in the sense that the grounds for the elimination of a connective, as given by its elimination rule, must not outstrip the grounds for introducing the connective, as given by its introduction rule. Some inferentialists have defended harmony while rejecting conservativity (e.g. see [Rea00]). Inferentialism is typically restricted to the view that the meaning of some proper *fragment* of our vocabulary is given completely by inferential rules. This is sometimes referred to as *moderate inferentialism*. One might wish to extend inferentialism to language as a whole, but it is unlikely that an advocate of any such program would endorse harmony as a meaning-theoretic constraint since it is far too constraining.

Should we require of empirical negation that it be governed by harmonious introduction-elimination rules? If empirical negation is non-logical and we are moderate inferentialists, then the answer is clearly "No". On the other hand, if empirical negation is deemed logical on grounds of topic neutrality, then the familiar intuitionistic negation will not be. For recall that empirical negation is an operation on warrants of which proofs are a special case, so in this sense it *generalizes* intuitionistic negation since it applies to a broader class of propositions, viz. the mathematical and empirical. Now it is unlikely that any inferentialist will accept this brief argument denying the logicality of intuitionistic negation on grounds of topic neutrality in which case they must hold that, while topic neutrality may be a sufficient condition on logicality, it cannot be necessary. According to an inferentialist, being characterizable by rules satisfying certain proof-theoretic constraints (e.g. harmony, purity, etc.) will be necessary and sufficient for logicality. But whatever conditions end up being necessary and sufficient for logicality, it is dubious that empirical negation will turn out logical on those grounds in which case the issue of harmony is irrelevant to empirical negation.[8]

This is not to say empirical negation fails to satisfy proof-theoretic constraints necessary and sufficient for logicality. Indeed one might hold that the burden of proof lies on the opponent to show that empirical negation *cannot* be given harmonious rules, an extremely strong claim to establish. But it is also debatable whether the rules given in section 3.5 for empirical negation are not harmonious. Read [Rea08] argues that labeled natural deduction systems for a variety of normal modal logics provide a way of furnishing harmonious rules for intensional connectives and if he is right then empirical negation, and a

---

[8]From a semantic perspective, empirical negation likely will not be counted logical either. For example, consider permutation invariance as a criteria of logicality. On the semantics of section 3.3 empirical negation is not permutation invariant, but neither is intuitionistic negation.

good deal of other connectives including non-logical ones, satisfy the harmony constraint.[9]

Heredity

Meyer and Routley have shown that relevant logic, once touted as a fierce rival to classical logic, can be viewed rather as an extension to classical logic with additional intensional (relevant) implication and (De Morgan) negation operators. One need only collapse the ordering $\leq$ on states, defined by

$$a \leq b \text{ iff } R0ab$$

for 0 the "base" state, by requiring that $a \leq b$ iff $a = b$. Collapsing the ordering to identity is required in order to ensure that the Heredity condition,

(Heredity)   If $M, a \models A$ and $a \leq b$ then $M, b \models A$,

is preserved when defining boolean negation $\sim$ by

(3.1) $M, a \models \sim A \text{ iff } M, a \not\models A.$[10]

    The formal semantics of relevant logic has taken a fair amount of criticism from its inception for not having an adequate informal interpretation and the collapse of the ordering that allows for boolean negation worsens that criticism.[11] For the ordering is supposed to represent an intuitive notion corresponding to something like information gathering, so that $a \leq b$ just in case b is an informational extension of a. The restriction requiring that $a \leq b$ iff $a = b$ claims that states have no proper extensions and embodies the unfounded Leibnizian optimism that each state is itself the best (in terms of informational content) amongst all possible ones.

    The point is that states are not perfect—i.e. complete and consistent. Moreover propositions, construed as sets of states, are not just any such sets: they are the hereditary ones. Now it is clear that boolean negation defined according to (3.1) is not an operation on propositions in the sense that the class of propositions is closed under that operation. For there are states a and b with $a \leq b$ such that a does not support a proposition A, and hence supports its boolean negation, while b supports A. But then the boolean negation of A is not preserved $\leq$-upward, and so the boolean negation of A is not a hereditary

---

[9]Read's approach is controversial for two reasons. The first is that while the rules are harmonious, the proof language makes explicit use of an "accessibility" relation and formula labels. Indeed, there are rules specifically for the accessibility relation. The question, then, is whether such proof languages presume a prior understanding of the relational semantics of modal logic and are thereby anti-inferentialist in spirit, or whether they undermine the inferentialist programme of *completely* determining the meaning of connectives by *just* the introduction (and possibly elimination) rules for the connective, and not partly by rules governing other connectives or (accessibility) relations. It is unclear whether labeled systems could be motivated along inferentialist lines independently of the semantic conception from which such systems arose.

[10]The literature on relevant logic refers to the negation defined by (3.1) as 'boolean' rather than 'classical', so I have stuck with this terminology. In a sense boolean negation is a natural characterization of classical negation for relational semantics, but I don't think it is the only natural one. Fixed negation seems to me to be just as natural.

[11]B. J. Copeland has been a main critic of relevantist semantics, charging them with being merely "pure" rather than "applied". See e.g. [Cop79], [Cop83] and [Cop86].

set (read 'proposition'). As boolean negation is just set-theoretic complementation in disguise, another way of putting this is to say that complementation is not an operation on propositions.

Notice that the same objection may be leveled against classical negation by any intuitionist who takes seriously a semantics on which boolean negation fails to be an operation on propositions intuitionistically conceived. In Kripke semantics for intuitionistic logic, propositions, being $\leq$-closed (for $\leq$ a preorder), are not closed under complementation, i.e. boolean negation. Now the intuitionist might attempt the same move as the relevantist by collapsing the preorder to identity (obtaining the class of "sheer reflexive" frames), but then the resulting logic is classical. So this move, available to the relevantist, is not available to the intuitionist. The intuitionist must reject boolean negation as a genuine propositional operation.

DeVidi and Solomon [DS06] work around this problem by introducing an empirical negation, defined in the context of Kripke semantics, that satisfies Heredity. Roughly the idea is this. The intuitionistic negation $\neg A$ of A is supported at a state iff A is not supported at any ($\leq$-) later state. We might think of some subset of states as having a property special to empirical negation, call this property being "actualized" (as put in [DS06]) at some time, and define negation relative to this property. Then we might think the empirical negation $\sim A$ of A is supported at a state just in case it is not supported at any later state *that is actualized*. To be precise, let $M = (W, A, \leq, V)$ be a usual Kripke model for intuitionistic logic with $A \subseteq W$ an additional set of actualized states. Define the truth conditions for $\sim$ by

- $M, a \models \sim A$ iff $\forall b$, if $a \leq b$ and $b \in A$ then $b \not\models A$.

It is easy to see that Heredity holds for the language extended with $\sim$, that is, $\sim$ is a genuine propositional operation.

There are two serious problems with this "negation". The first is that it is too weak. Almost no properties thought *characteristic* of negation hold for it. In particular, the law of excluded middle, a property we should think holds for empirical negation if we follow Dummett's remark quoted earlier, fails as does the law of non-contradiction and various directions of the De Morgan equivalences (e.g. $\sim(A \wedge B) \models \sim A \vee \sim B$). Second, it does not get things right at the level of models for there will be states supporting $\sim A$ for every A, i.e. states at which trivialism holds, a thesis rejected by any constructivist. This will be the case, e.g., when no later state is actualized in which case the truth conditions for $\sim A$ will be vacuously satisfied for arbitrary A.

For these reasons the empirical negation to be defined in section 3.3 is better motivated from a philosophical point of view and it also fairs much better in terms of getting correct the inferences that ought to intuitively hold of empirical negation.

Implicit vs explicit information

Relevantists (e.g. Greg Restall) have argued that there is a "difference between claims about states, and claims supported by states" [Res99, p. 71]. A state a's failing to support a proposition A should not imply that a supports some other proposition $\rightarrow A$ expressing a's lack of support for A—for this other proposition is *about a*. That is, it should not be assumed generally that a state supports

all the information about itself. An intuitionist who even only weakly endorses Kripke semantics could file the same charge against empirical negation if introducing it into the semantics requires making the unjustified identification of implicit and explicit information.

Of course some states might support all the information about themselves, but to say that all do is to make an unfair assumption about states. It would, however, be just as unfair to assume that states *never* support some or all of the information about themselves. Moreover, why not think that every model ought to possess at least one such state intended to represent a privileged state of the model, for example, the *present state of available evidence (or warrant)*? Such states will be complete in the sense that every statement will be either warranted or not.

Given Dummett's remark claims of the form '*A* is warranted' are in principle decidable, we should think that the present state of available evidence either warrants or fails to warrant the assertion of a given proposition. That is, the present state of available evidence is complete relative to empirical negation $\sim$ in the sense that $A \vee \sim A$ holds good at it. In modeling empirical negation, then, we should include at least one state representing the present state of available evidence for any claim of the form $\sim A$ asserts that, at present, A lacks sufficient evidence to be warranted. Such a state will form a classical model in the sense that the set of sentences in the language involving only the connectives $\wedge$ and $\sim$ will be maximally classically-consistent.

3.3   The semantics of fixed negation

Since we will be interpreting a connective $\sim$ to be read 'There is insufficient evidence at present[12] to warrant the proposition that. . . ' ($\neg$ will continue to be read 'It is refutable that. . . ') we must distinguish one state in each model as the present moment. Such a state will be further singled out by the truth conditions given to our empirical negation $\sim$. These informal ideas motivate the semantical clauses for the connectives of our language, to be specified in a moment, the only novel one being that for $\sim$.

Our language $\mathcal{L}^\sim$ is a usual language $\mathcal{L}$ for **IPC** (Intuitionistic Propositional Logic) augmented with the logical symbol $\sim$.[13] Let $F = (W, \leq)$ be a Kripke frame for the language $\mathcal{L}$ of **IPC**. That is, W is a non-empty set of states and $\leq$ is a partial order (reflexive, transitive and anti-symmetric relation) on W. We denote by $W\uparrow$ the set of all *upsets* in W, i.e. sets X s.t. $y \in X$ whenever $x \in X$ and $x \leq y$. An $\mathcal{L}^\sim$-model M is a tuple $(F, @, V)$ where $@ \in W$ is a distinguished element representing the state of available evidence in M (other

---

[12]The usual intensional semantics for temporal indexicals such as 'at present' (i.e. 'now') treats such expressions as non-indexical propositional operators, and this is how we have chosen to treat them here. (The same is true of non-temporal expressions such as 'actually' which is taken, by e.g. Lewis, to be indexical.) I do not find any concern, from a formal point of view, in semantically treating indexical expressions non-indexically. That is, we should not think that the formal semantics should be such that $\sim A$ is true at a state a just in case A is not true at a because $\sim$, being an indexical, should refer back to the state of evaluation. In other words, being an indexical does not imply being treated as an operator that cannot cause a state-shift away from the state of evaluation.

[13]For definiteness, the set of logical symbols is $\{\wedge, \vee, \rightarrow, \sim, \bot\}$, and *Prop* is a denumerable set of propositional letters (constants) whose members we denote by p, q, etc. We define $\neg A := A \rightarrow \bot$.

modally inequivalent states representing substates or states containing evidence unavailable in M) and $V : Prop \to W\uparrow$ is a propositional valuation assigning propositions (i.e. upsets in W) to propositional letters.[14]

We arrive at the following truth conditions (writing $M, a \models A$ to mean A is supported by a state a in the model M):

- $M, a \models A \wedge B$ iff $M, a \models A$ and $M, a \models B$;

- $M, a \models A \vee B$ iff $M, a \models A$ or $M, a \models B$;

- $M, a \not\models \perp$ (i.e. it is never the case that $M, a \models \perp$);

- $M, a \models A \to B$ iff $\forall b \geq a(M, b \models A \Rightarrow M, b \models B)$;

- $M, a \models {\sim}A$ iff $M, @ \not\models A$.

It is easy to verify by induction on formula complexity that (Heredity),

$$M, a \models A \text{ and } a \leq b \text{ imply } M, b \models A,$$

holds, as in the case of **IPC**, since $M, a \models {\sim}A$ iff for all $b \in W$, $M, b \models {\sim}A$. One may immediately notice the similarity of $\sim$ to the satisfaction operators of hybrid logic, the "actually" operator, and the "now" operator of temporal logics.

Truth in a model is truth at $@$, *validity on a frame* is truth in every model based on that frame, and *validity* "simpliciter" is validity on every frame. We say that a sentence A is an $\mathcal{L}^{\sim}$-consequence of set Γ of sentences, in symbols $\Gamma \models A$, iff for every model M, A is true in M whenever every member of Γ is true in M. We denote the set of valid consequences $\{(\Gamma, A) : \Gamma \models A\}$ by **IPC**$^{\sim}$.

We might have required that all models be rooted, i.e. that there be a minimum with respect to the partial order, or, perhaps more in line with common informal interpretations of intuitionistic models, that the models be tree-like. In any case, one obtains the same semantic consequence relation so we have chosen not to restrict our models in any of these two ways. It is interesting to note, however, that there are differences immaterial to consequence: e.g. if $@$ is always the root then $A \vee {\sim}A$ holds at *every* point of every model, whereas this is not true otherwise.

There are two notions of truth in a model we may distinguish: (i) *actual* truth which is truth at the distinguished element, and (ii) *global* truth which is truth common to all points in the model. A sentence is actually true in $M = (W, \leq, @, V)$ just in case it is true at $@$, and a sentence is globally true in M just case it is true at every state $a \in W$ of M. These notions of truth are obviously quite distinct. We may also distinguish between two types of consequence: *local* and *actual* consequence. A sentence A is a local consequence of a set Γ of sentences just in case for each model $M = (W, \leq, @, V)$ and every b in M, if every member of Γ is true at b then so is A. And A is an actual

---

[14]We may read $\leq$ in a number of different ways. My preferred is in terms of informational containment so that $a \leq b$ reads 'b contains all of the information contained in a'. Even though $@$ is to be thought of as the present state of evidence, we need not read $\leq$ temporally, e.g. by reading $a \leq b$ as 'b is a state of evidence arrived at temporally later than a which contains all the information contained in a' for we might just think of each state in the model as representing a *possible* present state of evidence.

consequence of Γ just in case A is true at @ whenever every member of Γ is. For the base intuitionistic language $\mathcal{L}$, the distinction between actual and local truth is one without a difference.[15] However, this is not true for $\mathcal{L}^\sim$ since now the distinguished element plays a significant role, where before it did not.

I wish to quickly settle a possible objection with the proposed semantics. Hossack writes

> [i]f a sign is to be regarded a negation of p at all, it must be used in such a way as to be incompatible with the assertion that p. The semantic rule has to hold, which Dummett calls Exclusion, that p and its negation cannot both be true. [Hos90, p. 216].

One possible objection to the semantics for negation introduced here is that we have states at which both A and $\sim A$ are true (supported) in violation of Hossack's necessary condition on being a negation operation. That's so, but never is it the case, in accordance with Hossack, that A and $\sim A$ both be *true*, for truth (relative to a model) as we have defined it is truth at @.

One interesting application of $\sim$ is that it may also be seen as providing an alternative, but inequivalent, characterization of boolean negation in the setting of relevant logic. If we confine ourselves to the simplified semantics for $\mathbf{R}$ then we may introduce $\sim$ by letting @ be the sole base world. Valid consequence on the simplified semantics for $\mathbf{R}$ is already defined in the same way we have defined valid consequence for $\mathbf{IPC}^\sim$, viz. in terms of truth preservation at @. It would be interesting to compare $\sim$ and boolean negation in the setting of relevant logic, a curiosity I mention only to set aside.

Before moving onto the next section, it is worth discussing an issue that has been raised concerning monotonicity and warrant. Tomassi [Tom06] states:

> ...non-monotonicity naturally suggests itself as a logical characteristic definitive of [defeasible warrant]. At least, it is difficult to see how warrant could genuinely be defeasible if it is not the case that further information could be obtained to defeat an assertion so warranted. To allow that defeating information can turn up however is precisely to allow the possibility that while X defeasibly warrants Z, the conjunction of X and Y might warrant the negation of Z...there is no obvious way of capturing the non-monotonic character of defeasible warrant within the framework provided by Kripke semantics for intuitionist logic. [Tom06, p.37-38]

While I agree that defeasible warrant introduces an element of non-monotonicity, I do not think it manifests at the level of logical consequence. Defeasible warrant is non-monotonic over the progression of time relative to some partial ordering on states of evidence. That is, if a and b are states of evidence such that $a \leq b$, where $\leq$ is some *monotonic* ordering of strength of evidence, and A is warranted at a, then A is warranted at b. What we do not have is that if

---

[15]A proof showing local validity implies actual validity is easy; the converse is not much harder. $\mathcal{L}$-models are $\mathcal{L}^\sim$-models and the truth conditions for $\mathcal{L}$-sentences are the same as those for $\mathcal{L}^\sim$-sentences restricted to $\mathcal{L}$, where $\mathcal{L} \subset \mathcal{L}^\sim$. For reductio, suppose A is an actual but not local consequence of Γ. Then there is a model $M = (W, \leq, @, V)$ and $a \in M$ s.t. $M \models \Gamma$ but $M, a \not\models A$. But then the model $M' = (W, \leq, @', V)$, where $@' = a$, is s.t. $M', @' \models \Gamma$ and $M', @ \not\models A$, which contradicts our supposition that A is an actual consequence of Γ.

A is warranted at some time t and $t < u$ then A is warranted at u. One has to distinguish between the orderings over which empirical truth is preserved. A statement A is warranted, in our sense, at a particular time t if and only if *all* the available evidence at t warrants A; i.e. statements are warranted relative to bodies of evidence. We should not want statements to be warranted relative to sets of statements, as they would be if reasoning with a defeasible warrant operator were non-monotonic. If we know that $A, \vdash B$ but $A, C \nvdash B$, and both A and C are warranted, then B is not warranted relative to $\{A\}$— it is simply not warranted. Ideally one should distinguish in the model two separate orderings, one $\leq$ over evidence and another $\preceq$ over temporal states, and introduce a future or past temporal operator $\square$ into the language governed by $\preceq$ such that A may hold according to some body of evidence at a moment t even though $\square A$ may not hold according to that same body of evidence and moment because A fails at the relevant $\preceq$-successors of t.

3.4 Validities and invalidities

The following lists some validities and inferences involving $\sim$ and shows a significant number of similarities between empirical and classical negation, most notably the holding by the former of all of the De Morgan equivalences, DNE, LEM and LNC, rule-form (EFQ), and classical reductio ad absurdum (RAA).

$$
\begin{array}{ll}
A \vee \sim A & \sim\sim A \rightarrow A \\
(\sim A \rightarrow A) \rightarrow A & \neg A \rightarrow \sim A \\
\sim(A \wedge B) \leftrightarrow (\sim A \vee \sim B) & \sim(A \vee B) \leftrightarrow (\sim A \wedge \sim B) \\
\sim(A \rightarrow B) \rightarrow \sim B & (A \wedge \sim B) \rightarrow \sim(A \rightarrow B) \\
\neg\sim A \rightarrow A & \sim(A \wedge \sim A) \\
A \wedge \sim A \models B & A \rightarrow B \models \sim B \rightarrow \sim A
\end{array}
$$

The following lists some notable exclusions to the above list, in particular all of the $\rightarrow$-forms of contraposition.

$$
\begin{array}{ll}
(\sim A \rightarrow \sim B) \rightarrow (B \rightarrow A) & (A \rightarrow B) \rightarrow (\sim B \rightarrow \sim A) \\
(A \wedge \sim A) \rightarrow B & \sim(A \rightarrow B) \rightarrow (A \wedge \sim B) \\
\neg(A \wedge \sim A) & \sim A \rightarrow \neg A
\end{array}
$$

Nearly all of the rule-forms of contraposition fail as well, a feature familiar to relevant logicians who have observed these sorts of failures for boolean negation in the setting of classical relevant logic.

It is worth noting that **IPC**$^\sim$ is not closed under substitution of provable equivalents, though it is obviously closed under uniform substitution. A logic is closed under substitution of provable equivalents when $\Gamma \vdash A$, $B \vdash C$ and $C \vdash B$ imply $\Gamma \vdash A(B/C)$, where $A(B/C)$ is the result of replacing some occurrences of B in A with occurrences of C. A logic is closed under uniform substitution when $\Gamma \vdash A$ implies $\Gamma \vdash A(p/q)$ for p, q atoms. For example $p \wedge \sim p \vdash p \wedge \neg p$ and $p \wedge \neg p \vdash p \wedge \sim p$, but $\vdash (p \wedge \neg p) \rightarrow \bot$ while $\nvdash (p \wedge \sim p) \rightarrow \bot$. This is not a particularly uncommon phenomenon. For example, compare logics with "actually" operators[16], temporal logics with "now" operators, and

---

[16]For instance, see [Hod84] and his notion of "weak" consequence according to which A and $@A$, the latter read 'actually A', are provably equivalent but not intersubstitutable *salva veritate*.

paraconsistent logics with non-truth-functional negations such as those of de Costa 1974, each of which fails to be closed under substitution of provable equivalents.

More importantly, $A \models B$ is weaker than $\models A \rightarrow B$ since the latter implies the former but not conversely—in other words, conditional proof fails. This explains, e.g., the $\rightarrow$-form failure of EFQ; $\rightarrow$ may take us to states at which both A and $\sim A$ hold, and hence states where an arbitrary A (e.g. $\bot$) need not (or in the case of the example, must not) hold. Often it is the implication from $\models A \rightarrow B$ to $A \models B$ that fails. For instance, this is the case with Tennant's relevant logics **CR** and **IR** (see [Ten97]). In particular, $\models (A \wedge \neg A) \rightarrow B$ holds but $A, \neg A \not\models B$ which is quite the opposite from **IPC**$^\sim$.

Remarks

Given the failure of conditional proof, **IPC**$^\sim$ is not an axiomatic extension of **IPC** in the sense that the former is obtainable from the latter by the addition of axioms. For any axiomatic extension of a system satisfying conditional proof, i.e. the deduction theorem in the context of axiomatic systems, satisfies conditional proof. However, **IPC**$^\sim$ is a conservative extension of **IPC**. This is easy to see semantically, a self-evidence not usually exuded on the proof-theoretic side. For if A does not contain $\sim$ then it receives the truth conditions it has in **IPC**-models and so if it were **IPC**$^\sim$-valid then it would have already been **IPC**-valid. As a final remark we note that **IPC**$^\sim$ is Tarskian in that it satisfies the structural rules of Reflexivity, Thinning and Cut.

Digression: It should not be surprising that the disjunction property, if $\Gamma \vdash A \vee B$ then $\Gamma \vdash A$ or $\Gamma \vdash B$, fails for **IPC**$^\sim$. This sort of failure is notorious for extensions of **IPC**. (So notorious in fact that it was conjectured by Łukasiewicz that **IPC** is the strongest intermediate logic to have the disjunction property.[17] Kreisel and Putnam proved this conjecture false by showing that the system obtained by adding $(\neg A \rightarrow (B \vee C)) \rightarrow ((\neg A \rightarrow B) \vee (\neg A \rightarrow C))$ to **IPC** has the disjunction property.) However, given our intended interpretation of the language it is not hard to justify the failure of conditional proof and the disjunction property. The justification for the failure of the latter rests on the assumption that it is decidable whether or not a proposition is supported by the evidence available at present in which case we expect $\vdash A \vee \sim A$ for arbitrary A without either of $\vdash A$ or $\vdash \sim A$ holding. So we focus our attention on the failure of conditional proof. **End of digression**.

There are two things to say about the failure of conditional proof. The first is that if we are to take seriously Kripke semantics for intuitionistic logic and if adding empirical negation is semantically acceptable then there is reason to reject conditional proof. After all, it falls out of an acceptable semantics for the language. Now most intuitionists do not find Kripke semantics acceptable primarily on the grounds that it is usually couched within a classical metatheory. But at the propositional level, this should not matter since **IPC** and its extension **IPC**$^\sim$ to fixed negation are complete with respect to a recursive set of finitely recursive models (i.e. models with a finite set of states and a recursive accessibility relation and valuation defined on that set). Indeed, if $\Gamma \not\vdash_{\textbf{IPC}} A$

---

[17]Technically, **IPC**$^\sim$ is not an intermediate logic since its language is not L but a proper extension of it. It might be more appropriate to call it a 'superintuitionistic' logic though some authors use these terms synonymously.

then, via a filtration technique, one can construct a finite countermodel from the set of subformulae of $\Gamma \cup \{A\}$ (assuming Γ is finite). Restricting the class of models so, the metatheory being classical or intuitionistic then come to the same thing.

While this fact does not generalize to quantificational theories, one may still furnish such theories with a constructive metatheory employing Kripke models even though certain results classically provable will not be provable in the constructive metatheory. This will not be problematic for a constructivist who finds something appealing with an informal interpretation of Kripke structures for intuitionistic logic and who wishes to make use of those structures in providing a semantics for her constructive language while sticking to a constructively acceptable metatheory. I shall therefore suppose that there are constructively acceptable versions of Kripke semantics for intuitionistic logic.

Despite the failure of conditional proof for $\rightarrow$, we may introduce other implication connectives satisfying conditional proof. One natural way to do this is to define A as being a consequence from Γ, where Γ is finite, just when $\models \bigwedge \Gamma \rightarrow A$. (Alternatively we could allow infinite sets of premises and infinitary conjunction, but this is unnecessary given compactness for $\mathbf{IPC}^{\sim}$.) Under this definition of consequence, consequence and implication coincide but the definition looks *ad hoc*. Moreover we lose important theorems that were motivated by our original concerns. For example EFQ for $\sim$ fails (since now $\bot$ would not be a consequence of $A \wedge \sim A$) as do a number of other important consequences involving the conditional that were valid under the original definition. So the revised definition of consequence does not speak well to the informal interpretation we had originally given to the language.

A more interesting conditional that satisfies conditional proof exploits the fact that (the theory of) @ is essentially a classical model with an additional intensional connective $\rightarrow$. Notice that *all* the classical connectives are definable at @ in a straightforward way, e.g. $A \supset B := \sim A \rightarrow \sim\sim B$ (or equivalently $\sim A \vee \sim\sim B$) says that *at an arbitrary point*, either A fails to be supported at @ or B is supported there, i.e. that A materially implies B at @. Then $\Gamma \vdash A$ (under the original definition of $\vdash$) iff $\vdash \bigwedge \Gamma \supset A$ for finite Γ. One with a strong liking for conditional proof might then think of material implication as *the* correct notion of implication for empirical discourse.

3.5 Tableaux

Several complications present themselves on the side of proof theory. The first, which we have already encountered, is that conditional proof fails. This means that in the setting of natural deduction, a $\rightarrow$-introduction rule would, in our case, turn out quite cumbersome. (This of course is not true for any system for which conditional proof fails.) The second, more troubling, difficulty is that since $A, \sim A \models \bot$ which, in a natural deduction setting, we might call $\sim E$, and $\neg A \models \sim A$, it follows that any of the usual presentations of natural deduction would generate a consequence relation satisfying both $\neg A \vdash \sim A$ and $\sim A \vdash \neg A$ (the latter following by $\sim E$ and $\neg I$) which is clearly bad (in respect of the latter); for $\sim$ is supposed to be strictly weaker than $\neg$. Exactly the same problem emerges when adding classical negation to intuitionistic logic in the usual systems of natural deduction.

For this reason we have chosen to give a *signed and labelled* tableau system for $\mathbf{IPC}^{\sim}$. The system is standard for $\mathbf{IPC}$ for this style of tableau system and adds two additional rules for sentences of the form $S{\sim}A$ where S is a sign in $\{\mathrm{T},\mathrm{F}\}$ representing the truth or falsity of a sentence at a state (e.g. '$\mathrm{T}A\wedge B, i$' expresses that $A\wedge B$ is true at state i). For tableaux there are two types of sentences.

Definition 3.5.1.

- A *signed labelled sentence* has the form SA, i for $S \in \{\mathrm{T},\mathrm{F}\}$, A an $\mathbf{IPC}^{\sim}$-sentence and $i \in \{@\} \cup \mathbb{N} = Labels$ (the set of labels).

- A *relational sentence* has the form $i \leq j$ for $i, j \in Labels$.

A tableau *for an argument* $A_1, \ldots A_n \Rightarrow B$, with A_i the premises and B the conclusion, starts by listing the $\mathrm{T}A_i, @$ in their natural order proceeded by $\mathrm{F}B, @$. More precisely we have the following definition of a tableau.

Definition 3.5.2 (Tableaux)**.**

- A *tableau* $\mathcal{T} = (X, \leq)$ is a partially ordered, finite set of sentences (labelled or relational) with a minimum (root) such that for each $\alpha \in \mathcal{T}$ the set of predecessors of α under $\leq$ forms a linear order. If $\leq$ is a partial order on a set X then α is a minimum in X w.r.t. $\leq$ iff for all $\beta \in X$, $\alpha \leq \beta$.

- A *branch* of a tableau $\mathcal{T}$ is a maximal chain in $\mathcal{T}$.

Definition 3.5.3 (Closure)**.**

- A *branch closes* iff there is an $i \in Labels$ and sentence A such that both $\mathrm{T}A, i$ and $\mathrm{F}A, i$ occur on the branch.

- A *tableau closes* iff each of its branches closes.

- A *branch (tree) is open* if it is not closed.

A closed tableau for an argument is a proof of that argument. An argument $A_1, \ldots A_n \Rightarrow B$ is provable just in case there exists a closed tableau for it, in which case we write $A_1, \ldots A_n \vdash_{\mathbf{IPC}^{\sim}} B$, sometimes suppressing the subscript when it is clear.

We have the following rules, a pair for each connective $\otimes$ indicating how to decompose sentences preceded by T or by an F whose main connective is $\otimes$.

$$
\begin{array}{cccccc}
\mathrm{T}{\sim}A, i & \mathrm{F}{\sim}A, i & \mathrm{T}\neg A, i & \mathrm{F}\neg A, i & \mathrm{T}A \to B, i & \mathrm{F}A \to B, i \\
| & | & irj & | & irj & | \\
\mathrm{F}A, @ & \mathrm{T}A, @ & | & irj & \diagup\;\diagdown & irj \\
& & \mathrm{F}A, j & \mathrm{T}A, j & \mathrm{F}A, j \quad \mathrm{T}B, j & \mathrm{T}A, j \\
& & & (\text{new } j) & & \mathrm{F}B, j \\
& & & & & (\text{new } j)
\end{array}
$$

$$
\begin{array}{cccccc}
\mathrm{T}A \wedge B, i & \mathrm{F}A \wedge B, i & \mathrm{T}A \vee B, i & \mathrm{F}A \vee B, i & irj & SA, i \\
| & \diagup\;\diagdown & \diagup\;\diagdown & | & jrk & | \\
\mathrm{T}A, i & \mathrm{F}A, i \quad \mathrm{F}B, i & \mathrm{T}A, i \quad \mathrm{T}B, i & \mathrm{F}A, i & | & iri \\
\mathrm{T}B, i & & & \mathrm{F}B, i & irk &
\end{array}
$$

The parenthetical '(new j)' indicates that the label j must be new to the branch; i.e. it must not occur as the label of any formula on the branch besides those introduced by the rule. In the second to last rule of the second row, governing the transitivity of r, jrk does not need to immediately follow irj on the branch. All that is required is that each occur on the branch. The same may be said about the conclusions of the other rules, though it need not be. For definiteness we understand the other rules to be read in the same light.

3.6 Completeness

The proof of completeness for the tableau system is given according to the usual recipe. We confine ourselves to *systematic tableaux*, which means that the rules are to be taken as "musts" rather than "cans": if a rule can be applied, it must be applied. A systematic tableau is one in which every branch β is *downward saturated* (defined below). The goal then is to show that every open branch of a systematic tableau forms a Hintikka set (defined below), and that every Hintikka set is satisfiable. Completeness then follows almost immediately.

Definition 3.6.1. A branch β is *downward saturated* provided the following hold.

1. If $\mathrm{T}A \wedge B, i$ occurs on β then $\mathrm{T}A, i$ and $\mathrm{T}B, i$ occur on β.

2. If $\mathrm{T}A \vee B, i$ occurs on β then either $\mathrm{T}A, i$ occurs on β or $\mathrm{T}B, i$ occurs on β.

3. If $\mathrm{T}\neg A, i$ and irj occurs on β then $\mathrm{F}A, j$ occurs on β.

4. If $\mathrm{T}A \rightarrow B, i$ and irj occurs on β then either $\mathrm{F}A, j$ occurs on β or $\mathrm{T}A, j$ occurs on β.

5. If $\mathrm{T}{\sim}A, i$ occurs on β then $\mathrm{F}A, @$ occurs on β.

6. If $\mathrm{F}A \wedge B, i$ occurs on β then either $\mathrm{F}A, i$ occurs on β or $\mathrm{F}B, i$ occurs on β.

7. If $\mathrm{F}A \vee B, i$ occurs on β then both $\mathrm{F}A, i$ and $\mathrm{F}B, i$ occur on β.

8. If $\mathrm{F}\neg A, i$ occurs on β then for some j, irj and $\mathrm{T}A, j$ occur on β.

9. If $\mathrm{F}A \rightarrow B, i$ occurs on β then for some j, irj, $\mathrm{T}A, j$ and $\mathrm{F}B, j$ occur on β.

10. If $\mathrm{F}{\sim}A, i$ occurs on β then $\mathrm{T}A, @$ occurs on β.

11. For all i occurring on β, iri occurs on β.

12. If irj and jrk occur on β then irk occurs on β.

Definition 3.6.2. A set H of formulas is a *Hintikka set* if it satisfies the conditions of Definition 3.6.1 by replacing 'occurs on' by 'is a member of' and 'β' by 'H', and the further condition that for no propositional letter p and $i \in Labels$ are both $\mathrm{T}p, i$ and $\mathrm{F}p, i$ in H.

Lemma 3.6.3. *Every open branch β forms a Hintikka set. i.e. the set of formulas consisting of the nodes of β is a Hintikka set.*

Proof. Immediate by Definitions 3.5.2 and 3.6.2. □

Lemma 3.6.4 (Model existence)**.** *Every Hintikka set H is satisfiable in an* **IPC**$^\sim$*-model $M^H = (W, \leq, @, V)$ defined as follows. Let label(H) be the set of labels in H, i.e. label$(H) = \{i : SA, i \in H \land S \in \{T, F\}\}$. Set*

- $W = label(H)$

- $i \leq j$ *iff* $irj \in H$

- $@ = @$

- $V(p) = \{i \in W : Tp, i \in H\}$.

Proof. We do two things. The first is to verify that M^H is indeed a model. The second is to show that the Fundamental Lemma 3.6.5 holds.

We claim that $\leq$ is a partial order. Clearly for all $i \in W$, $i \leq i$ by definition and the fact that for each $i \in label(H)$, $iri \in H$ by the "reflexivity" rule. Transitivity follows similarly by the definition of $\leq$ and the fact that $irj, jrk \in H$ only if $irk \in H$ by the "transitivity" rule and downward saturation. Anti-symmetry is vacuously satisfied since the rules never allow that $irj, jri \in H$ for $i \neq j$.

To finish verifying that the structure defined above is an **IPC**$^\sim$-model we have only left to show that for no $i \in W$ and $A \in L^\sim$ do we have $i \in V(A)$ and $i \in \overline{V(A)}$, where $\overline{X}$ is the W-complement of X. As the result follows trivially by the definition of V (as the reader may wish to verify), our verification that M^H is a model is concluded.

All that is left is proof of the following "Fundamental Lemma".

Lemma 3.6.5 (Fundamental Lemma)**.** *If $TA, i \in H$ then $M^H, i \models A$, and if $FA, i \in H$ then $M^H, i \not\models A$.*

Proof. We proceed by induction on formula complexity.

Base. *A* is an atomic letter p or $\bot$. The case for $\bot$ is trivial, so suppose $Tp, i \in H$. Then $i \in V(p)$, so $M^H, i \models p$.

Now suppose $Fp, i \in H$. Then $i \notin V(p)$, so $M^H, i \not\models p$.

Now assume the induction hypothesis (IH) holds for all formulas of complexity less than A's.

Case 1. *A* is $B \wedge C$. Suppose $TB \wedge C, i \in H$. Then $(TB, i), (TC, i) \in H$, so By IH $M^H, i \models B$ and $M^H, i \models C$. Hence $M^H, i \models B \wedge C$.

Now suppose $FB \wedge C, i \in H$. Then either $FB, i \in H$ or $FC, i \in H$. By IH either $M^H, i \not\models B$ or $M^H, i \not\models C$, whence $M^H, i \not\models B \vee C$.

Case 2. *A* is $B \vee C$. Suppose $TB \vee C, i \in H$. Then either $TB, i \in H$ or $TC, i \in H$. By IH either $M^H, i \models B$ or $M^H, i \models C$, whence $M^H, i \models B \vee C$.

Now suppose $FB \vee C, i \in H$. Then both $FB, i, C, i \in H$. By IH $M^H, i \not\models B$ and $M^H, i \not\models C$, whence $M^H, i \not\models B \vee C$.

Case 3. A is $\neg B$. Suppose $T\neg B, i \in H$. Then for all j s.t. $irj \in H$ and $FB, j \in H$. By IH, for all j s.t. $i \leq j$, $M^H, j \not\models B$. Hence $M^H, i \models \neg B$.

Now suppose $F\neg B, i \in H$. Then there is a j s.t. $irj \in H$ and $TB, j \in H$. Then $i \leq j$ and by IH $M^H, j \models B$. Hence $M^H, i \not\models \neg B$.

Case 4. A is $B \to C$. Suppose $TB \to C, i \in H$. Then for each j s.t. $irj \in H$, either $FB, j \in H$ or $TC, j \in H$. By IH either $M^H, j \not\models B$ or $M^H, j \models C$ for each $j \geq i$. Hence $M^H, i \models B \to C$.

Now suppose $FB \to C, i \in H$. Then there is a j s.t. $irj \in H$, $TB, j \in H$ and $FC, j \in H$. Thus $i \leq j$ and by IH $M^H, j \models B$ and $M^H, j \not\models C$. Hence $M^H, i \not\models B \to C$.

Case 5. A is $\sim B$. Suppose $T\sim B, i \in H$. Then $FB, @ \in H$, whence by IH $M^H, @ \not\models B$, so $M^H, i \models \sim B$.

Now suppose $F\sim B, i \in H$. Then $TB, @ \in H$, whence by IH $M^H, @ \models B$, so $M^H, i \not\models \sim B$.

$\square$

This concludes the proof that every Hintikka set is satisfiable. $\square$

Theorem 3.6.6 (Completeness)**.** *If* $\Gamma \vdash B$ *then* $\Gamma \models B$. *(Recall that* $\Gamma \models B$ *means that for every model M, if $M \models \Gamma$ then $M \models B$.)*

Proof. Suppose $\Gamma \not\vdash B$. Then there is a systematic tableau $\mathcal{T}$ with open branch β s.t. for each $\gamma \in \Gamma$, $T\gamma, @$ occurs on β and $FB, @$ occurs on β. Let H be the Hintikka set whose elements are the nodes of β. Then by the fundamental lemma, there is an **IPC**$^\sim$-model M^H s.t. $M^H, @ \models \Gamma$ and $M^H, @ \not\models B$, thus $\Gamma \not\models B$. Contraposing, if $\Gamma \models B$ then $\Gamma \vdash B$. $\square$

3.7 Final remarks

We have considered an operation $\sim$ such that the truth of $\sim A$ at an arbitrary point of evaluation is determined completely by the truth of A at a distinguished point representing all evidence available at present. When consequence is defined as preservation of truth in a pointed model, $\sim$ looks nearly classical. Moreover, it seems plausible that any semantics defined with respect to a class of models which distinguishes a point according to which truth in that model amounts to truth at that point, solicits the characterization of a family of connectives defined in terms of the distinguished point.

If verificationism is a theory about ordinary discourse, then any intuitively coherent connective used in such discourse ought to admit of a constructively acceptable semantics where *expressive adequacy* is a concern. A theory T is expressively adequate with respect to a connective ♣ picked out by natural language expressions if there is a connective $\otimes$ in the language of T which corresponds, in an intuitive sense given the $\vdash_T$-rules governing $\otimes$, to ♣. For instance, **IPC** is expressively adequate with respect to conjunction to which $\wedge$ corresponds. **IPC** is clearly not expressively adequate with respect to weak and empirical negations, just as certain theories of truth (of e.g. Kripke) are not expressively adequate with respect to "exclusion negation" (as opposed to

"choice negation"), the operation which takes an intermediate truth value to truth in a three-valued semantics such as Kleene's.

This expressive inadequacy sometimes brings with it predictable paradoxes of an impoverished language as noted e.g. in [Wil94], where Williamson generates a Fitch-like paradox for an intuitionistic language enriched with an additional warrant operator K read 'at some past, present or future time someone possesses a warrant to assert A'. The argument runs as follows. The claim that A will never be decided may be formalized as $\neg KA \wedge \neg K \neg A$. But $\neg KA$ implies $\neg A$ (shown below) which, together with $\neg KA \wedge \neg K \neg A$, yields the contradiction $\neg A \wedge \neg\neg A$. Here is the argument Williamson gives, a mild variation of Fitch's original, to show that $\neg KA$ implies $\neg A$.

1.	$A \wedge \neg KA$	supposition
2.	$K(A \wedge \neg KA)$	1 BHK interpretation
3.	$KA \wedge K \neg KA$	2 distribution
4.	$KA \wedge \neg KA$	3 axiom T
5.	$\neg(A \wedge \neg KA)$	1,4 $\neg$I
6.	$\neg KA \rightarrow \neg A$	5 intuitionistic propositional logic

If in line (1) we replace $\neg$ with $\sim$ we can obtain everything up to the penultimate line (even with intuitionistic negation out front), but we can not obtain the conclusion, line (6).[18]

---

[18] Actually the paradox is blocked in a Kripke semantics for **IPC** with a standard-going **S5**-like epistemic operator K since $A \wedge \neg KA$ will then be consistent.

Chapter 4

Consequence relations for speech acts and propositional attitudes

Abstract

Rationality applies to a broad range of speech acts and propositional attitudes. The problem is that such acts and attitudes may not be rationally constrained by any of the usual notions of consequence-as-truth-preservation (for some given understanding of truth, e.g. truth as provability). In this chapter I consider some other notions of consequence required in formulating rationality constraints on doubting and denial in both classical and non-classical settings. A complete labelled natural deduction system is given for a logic of doubt which coincides with Strong Kleene logic.

4.1   Introduction

A unary operator ∇ in a language $\mathcal{L}$ (also the concept ∇ is intended to represent, if there is one) is closed under an implication or consequence relation $\vdash$ (between sets of $\mathcal{L}$-formulae and $\mathcal{L}$-formulae[1]) when

(Closure) $\qquad\qquad\qquad \nabla A$ and $A \vdash B$ imply ∇B.

∇ is closed under converse implication when

(C-Closure) $\qquad\qquad\qquad \nabla B$ and $A \vdash B$ imply ∇A.

Typically one thinks of consequence in terms of truth preservation, so that for a set Γ of sentences, $\Gamma \vdash A$ precisely when, relative to some class $\mathcal{C}$ of

---

[1]Set-theoretic braces are omitted on the left of $\vdash$ according to the usual convention, and the comma denotes set-theoretic union.

models or interpretations for $\mathcal{L}$, A is true under any $M \in \mathcal{C}$ whenever each member of Γ is true under M. Some examples of operators closed under implication are epistemic (knowledge) and doxastic (belief) operators when the agents intended to be modeled are suitably idealized (e.g. they may be taken to be "perfectly rational"). When the implication relation is representable by an object-language operator $\Rightarrow$ (i.e. an operator of $\mathcal{L}$), we may state weaker forms of closure such as closure under *known* or *believed*, etc. consequence. For instance, when the agents are not perfectly rational (yet still ideal in a weaker sense) it is more plausible that an epistemic operator K be closed under known consequence rather than consequence unrestrictedly. That is, it is more plausible that whenever KA and $K(A \Rightarrow B)$ then KB, as this does not imply what has come to be called "logical omniscience"—that an agent know every consequence of whatever she knows.

It is quite familiar by now that a number of philosophically central intensional operators fail to be closed under consequence or even naturally weakened forms (e.g. known consequence), including deontic operators ("It ought to be the case that"), epistemic and doxastic ones and, as a paradigm case, negations. With regard to the latter, any such instance of closure might be seen as committing the fallacy of denying the antecedent, and so *no* instance ought to hold. This is a consequence of the fact that negation is closed under converse implication (and also supposing that it is not a trivial operator satisfying ∇A for all A) and the same is typically thought to hold for denial (taken as a speech act) and rejection (taken as a propositional attitude).[2]

However, there are a number of interesting cases where denial fails to be closed under converse implication because it fails to preserve falsity (though it will always preserve untruth assuming consequence preserves truth) and, moreover, falsity and denial are closely linked. (On some accounts it is not falsity but untruth that is closely linked with denial.) For example, the usual notions of superconsequence in supervaluationism fail to anti-preserve superfalsity, so if one thinks that whatever is superfalse ought to be denied, and perhaps nothing else, then she cannot formulate this ought as a (possibly weakened) closure condition along the lines of (Closure). In these cases consequence as truth-preservation plays little to no role in defining suitable constraints on rational denial and rejection, whereas consequence as falsity-preservation does. Thus one must distinguish two notions of consequence where the usual notion as truth-preservation, the notion *par excellence*, regresses into the shadows. In short, consequence as truth preservation is often not what we should or do care about, especially in the larger context of theories of speech acts and propositional attitudes.

There are many other cases where consequence as truth-preservation plays no role in closure constraints. Some of these involve attitudes such as doubting, wondering, entertaining, desiring, and so on. In these cases there may be no close connection between the attitude and either truth or falsity. Consider doubt. Should it be closed under consequence—or even any suitably weakened candidate—or its converse? The former is answered with a resounding "No". Mundane counterexamples abound: I may doubt A but I do not doubt $A \to A$ for any A. Notice that this has nothing to do with 'S doubts that...' being an

---

[2]e.g. see [Res05] and [Res08] for an endorsement that denial and rejection are closed under converse implication.

'intensional' operator which typically do not satisfy straightforward closure-under-consequence principles. Closure under converse implication is a more complicated matter, but on certain notions of doubt, e.g. as neither believing nor disbelieving, this fails too. There are, however, notions of consequence that do constrain doubt, though the semantic value which they preserve will be something like the intermediate value of a three-valued semantics such as Kleene's.[3]

In this chapter I wish to discuss the connections between three things: (i) the semantic values of a given semantic theory, (ii) consequence relations defined as preservation of those various semantic values, and (iii) plausible constraints placed on rationally performing certain speech acts (such as denials) or rationally possessing certain propositional attitudes (such as doubtings) according to (Closure). One might think that there is no tight connection between doubt and consequence, for one may doubt nearly anything they like (though recalling a lesson learned from Descartes, not everything), even if they believe it to be true. This intuition seems to be, in general, misguided as section 4.3 hopes to show.

The chapter proceeds as follows. In section 4.2 I discuss plausible philosophical motivations for rejecting the closure of denial under converse consequence-as-truth-preservation. One involves semantic paradoxes, another supervaluationism (which may arise from issues concerning vagueness or tense), and another involving intuitionism. In regard to the semantic paradoxes, I take up an argument which rejects any motivation the gap theorist might have for denying the liar on the purported grounds that the liar is gappy. This supports a general closure constraint on rational denial that mirrors analogous constraints placed on both assertion and doubt. Section 4.3 provides a constraint on doubt analogous to those for assertion and denial. A notion of doubt as agnosticism (and hence one that excludes disbelief as a sufficient condition for dubitability) is motivated by a close connection between (i) what may be asserted, denied and doubted in mathematical discourse and (ii) Kleene's three-valued semantics. A natural deduction system for consequence as indeterminacy preservation is given in section 4.6 and proved strongly complete.

4.2 Denial

It is typically thought (e.g. by Restall in [Res05] and [Res08]) that denial is constrained by implication in the following way mentioned above:

(4.1)   If $A \vdash B$ and B is deniable then so too is A,

where $\vdash$ denotes some relevant notion of consequence as truth-preservation. But this constraint holds only on a particularly *weak* notion of denial, a notion according to which A is deniable if it is not true. (Recall that something's being deniable does not mean it *ought* to be denied. It means it *may* be denied.) There is, however, a stronger sense of denial whereby knowing or having good reasons to believe a sentence is untrue does not always suffice

---

[3]I should mention that there are two interesting limit cases according to which doubt is closed under consequence as truth-preservation, viz. those of the trivialist and the "radical skeptic". The trivialist doubts nothing and the radical sceptic doubts everything, so each doubts whatever follows from whatever she doubts.

for its being deniable. Certain gap theorists will require that the denial of a statement expresses a strong attitude toward it, namely the rejection of that statement as being *false*. Since being untrue is not equivalent to being false for a gap theorist, the untruth of a statement will not always suffice for its being deniable. It would be indefensible for an intuitionist—and even a classical mathematician—to *deny* Goldbach's conjecture simply because it is presently undecided and perhaps even undecidable (recalling that for an intuitionist, truth is provability). She should remain agnostic instead, and deny the conjecture only if she has a *refutation* of it; for all she knows, it may be proved in the future. Since intuitionistic consequence does not anti-preserve refutability, denial cannot be constrained by provability-preserving consequence as required in (4.1).[4] We must instead distinguish two notions of consequence, consequence as falsity-preservation and consequence as truth-preservation, and formulate the constraint in terms of the former.[5]

For present purposes I wish to focus on the strong sense of denial because of its close association with consequence relations that do not anti-preserve some relevant notion of falsity, such as supervaluational consequence (as it is usually formulated as preservation of supertruth) and its failure to anti-preserve superfalsity. This is not to say that the weak notion of denial is not a sensible one or one that is often employed in ordinary language. To be sure, it is. One finds it featuring prominently in the literature on the liar paradox. Take the gap theorist's usual response to the liar sentence λ ('λ is untrue'), holding that it is neither true nor false. If they assert that λ is untrue then they are asserting what, by their own lights, is untrue. But (most claim) it is a platitude of assertibility that one ought to assert only what they take to be true (or at least warranted in some sense), and since the gap theorist does not take the liar to be true, she should not assert it. A typical response for the gap theorist is to deny that his utterance of λ is not an assertion of any kind but rather a denial, a denial that λ is true. Since λ is untrue and one ought to deny what is untrue—on our present weak account of denial—the gap theorist's denial of λ seems correct.

So it appears the gap theorist has a reason for endorsing weak deniability. But actually I think this is not the best response for her to make for two reasons. The first is that, on a weak account of assertion and denial, the two are no longer symmetric. There are two distinct grounds for the deniability of a statement, that it be true or false, and only one for its assertibility, that it be true. But what grounds are there for breaking the symmetry? Why should we not hold that the only grounds for deniability are falsity? The main reason for denying whatever is gappy is that the liar is gappy in a way that warrants its denial. But notice that not *all* gappy sentences are like this. According to certain theories such as Kripke's, the truthteller is also gappy (e.g. it is gappy

---

[4]A caveat: intuitionistic logic **IL** anti-preserves refutability when arguments are restricted to a single premise, but not otherwise. (Conjoining a finite number of premises doesn't help if we identify the denial of a conjunction of premises with the denial of at least one of the conjuncts, as we should.) For if $\vdash_{\textbf{IL}} \neg B$ and $A \vdash_{\textbf{IL}} B$ then $\vdash_{\textbf{IL}} \neg A$. Yet we do not have (like we do for classical logic), for Γ a set of sentences, that if $\Gamma \vdash_{\textbf{IL}} A$ and $\vdash_{\textbf{IL}} \neg A$ then there is a $\gamma \in \Gamma$ such that $\vdash_{\textbf{IL}} \neg\gamma$. That is, we don't have generally that if some set of premises implies a conclusion that is refutable then the premises are refutable too in the sense that at least one of its members is.

[5]There is a familiar and important distinction between a statement being provable and a statement having been proved. I have ignored them here but return to them in section 4.3.

in the minimal fixed-point), but it is not clear that one ought to deny it, for it is at least consistent (e.g. there are fixed-points in which it is true).[6] One might hold that only gappy paradoxical sentences should be denied, but this response seems *ad hoc*. In any case, I think the gap theorist has a much better response than to argue that an utterance of λ is really a denial that λ is true rather than an assertion that λ is not true.

Suppose A is gappy. Then on certain widely held views about the ontology of propositions (e.g. a Russellian one), A does not express a proposition, for every proposition is either true or false and gaps occur only at the level of sentences when they fail to express propositions. (I shall come to the gappy propositions view below, endorsed by Milleans (about proper names) who hold that some names fail to refer.) We may hold, quite sensibly, that A nevertheless is *meaningful* (supposing it is well-formed or grammatical). There is no reason to think that a sentence is meaningless if it fails to express a propositions or that propositions ought to do the work of "meanings".[7]

Now consider an alleged denial of our gappy A. Exactly what is one denying when she allegedly denies that A? It cannot be a proposition since A does not express one. But on most counts, denials are precisely denials of propositions and only derivatively of other things insofar as those things (e.g. sentences) express propositions.[8] A denial is a judgment about a *content*, and where there's no content, there's no denial. The same goes for assertion, belief, and so on (assuming propositions are the contents of these attitudes and acts). Thus when the gap theorist utters the liar, she is neither asserting nor denying anything. And to reemphasize, that doesn't mean he isn't saying anything meaningful.

This may seem counterintuitive but it is a simple consequence of two basic assumptions: (i) speech acts act on contents; (ii) gappy statements do not express contents. So if one holds that statements that fail to express propositions may be denied *in some sense*, she will need to spell out precisely the sense she has in mind. It will have to be a sense according to which at least one of (i) or (ii) is given up, and I do not think this leaves one with a plausible sense of denial.

One way of giving up (ii) and defending the view that gappy statements express contents is to hold a gappy view of (Russellian) propositions according to which what is expressed by a statement containing an irreferential expression (e.g. a singular term or even a predicate—though we will consider only gaps resulting form the former) is a gappy proposition. As an example, 'Pegasus is winged' expresses the gappy proposition $\langle \emptyset, \text{being winged} \rangle$, where $\emptyset$ represents the denotation of 'Pegasus'.[9]

---

[6]This assumes that there are no independent grounds for maintaining that the minimal fixed-point is *the* right model or has priority over others in modeling truth.

[7]Unfortunately, this has often been the view, but it seems philosophers are starting to tend away from it. For example, if one holds a truth-conditional theory of meaning then a sentence can be meaningful, in the sense of having truth conditions, without expressing a proposition. Of course, one need not endorse a truth-conditional theory of meaning to uphold the distinction between meaning and proposition.

[8]This view—that the object of denials are propositions—is endorsed by e.g. Frege, Priest in [Pri06], Smiley in [Smi96] and Parsons in [Par84].

[9]Other accounts simply put a blank or empty placeholder in for $\emptyset$. I prefer the use of $\emptyset$ since then propositions continue to be well-defined sets, on a par with non-gappy propositions. Braun in [Bra93] gives precisely this account as well.

Now let us consider the sentential liar, 'This sentence is untrue'. Is it gappy? We are supposing that the predicate 'is untrue' expresses a property of propositions, so if the liar is gappy it must be because its subject term 'This sentence' fails to refer. But it *does* refer. It refers to a sentence (type), viz. the liar. So the sentential liar does not express a gappy proposition. What about the propositional liar, 'The proposition expressed by this sentence is untrue'? If one believes that propositions just are set-theoretic objects or that propositions are best represented by set-theoretic objects and that there are no non-well-founded sets, then the propositional liar does not express a proposition. For if it did, it would have to be a non-well-founded set.

Indeed, no self-referential statement expresses a proposition on this account. The fact that self-referentiality is ruled-out at the propositional level, and on well-motivated grounds arising from considerations of the ontology of propositions, I take to be an advantage of the standard Russellian account of propositions. Those who feel there is some draw to Russellian propositions but also that self-referentiality ought not be barred at the level of propositions might wish to allow for non-well-founded sets, as is done in [BE87]. I think to endorse non-well-founded propositions is to confuse *representation* (e.g. cyclic graphs) with *the represented* (i.e. non-well-founded sets). I can see no philosophical grounds for thinking that non-well-founded sets *exist*.

One standard objection to holding that propositions are self-referential asks how we can account for the truth of 'This sentence is English' which I assume is true. That *sentence* is certainly self-referential but the proposition it expresses most certainly is not, for the proposition refers to a *sentence* rather than a proposition, let alone itself. How about the proposition expressed by 'This proposition is English'? It fails to express a proposition and even if it did, it would be false, since propositions on most accounts are not in a language such as English.

The only option left is to give up (i) and thereby hold that denials (and assertions and the like) act on something other than contents, such as statements. I find this rejoinder highly implausible. Statements are either meaningless, abstract syntactic expressions, and hence not the object of speech acts, or otherwise they are syntactic expressions endowed with things that are meaning-like or content-like. But these latter things just sound like propositions to me. Alternatively one might think that when one asserts A what she intends to express is something like 'A expresses a true proposition' and when she denies A what she intends to express is something like 'A either expresses a false proposition or no proposition at all'. But again, I see no reason to deny *every* statement that fails to express a proposition and so there is no reason to think that a denial of A amounts to expressing that A either expresses a false proposition or no proposition at all.

So it seems the best thing for the gap theorist to say concerning the liar— supposing she endorses a Russellian account of propositions and a well-founded hierarchy of sets—is that she neither asserts nor denies it, and that when she utters 'λ is not true' (where λ denotes that very sentence), she really asserts something—not whatever is the purported content of λ since there is none, but rather that λ fails to express a proposition and is hence untrue.

The consequence constraint and denial

Frege held that the denial of a statement amounts to no more than the assertion of its negation, or at least that the content of a denied statement is always equivalent to the content of an asserted one. He thus concluded that we may reduce the number of judgment primitives to but one—assertion. Of course by symmetry he could have argued the other way around—that the assertion of a statement can be defined in terms of the denial of its negation, but ignoring issues of conceptual priority, the point is simply that, out of the two notions, we need only one of them to do the required work.

The problem with Frege's equivalence thesis is that it fails according to any glut theory in which assertion and denial, or more plausibly assertibility and deniability, are mutually exclusive. Since a glut theorist both asserts A and $\neg A$ for certain A (e.g. the liar), she cannot maintain that the assertibility of $\neg A$ implies the deniability of A.

Actually she *could* maintain this since she is a glut theorist—it would constitute just another glut in her theory. But if she wants to express genuine disagreement with anyone she should not posit any gluts in her theory of assertion and denial. For if the two are not exclusive then I see no way for the glut theorist to express disagreement with others; asserting the negation will not do and neither will denying what the other asserts, and there is no point in positing other exclusive speech acts to do the work. It is worth noting that on a normative theory about what one *ought* to (rather than may) assert, Priest is fine with conflicting obligations (e.g. see [Pri06]), so what one ought to assert and what one ought to deny are *not* exclusive. Priest does, however, believe that the attitudes of acceptance and rejection are exclusive. If we think assertion merely expresses acceptance and denial merely expresses rejection, as likely most do (e.g. see [Smi96]), then *correctly* asserting and correctly denying will come to the same thing as accepting and rejecting, and so the two will be exclusive. From hereon I assume a notion of assertibility and deniability according to which the pair is exclusive, a property I take to be something of a platitude.

The equivalence thesis also fails for gap theories in which being gappy suffices for being deniable. If A is gappy then so too is $\neg A$ (assuming negation is choice negation), so both are deniable and hence neither is assertible. But not all gap theories allow denial to come so easily as from a mere lack of truth value, and those that do not will in general need to distinguish falsity-preserving consequence from truth-preserving consequence when formulating a constraint on deniability along the lines sketched above.

4.3   Doubt

Jim and Max are having a conversation about their plans to attend the theatre tonight:

Jim: I doubt Sarah will be at the theatre tonight.

Max: But Paula won't be there unless Sarah is.

Max: So you must doubt that Paula will be there too.

The inference made by Max seems *prima facie* valid. Indeed it looks like an instance of *modus tollens* where 'It is doubted that. . .' replaces negation, 'It is not the case that. . .' The reason the argument looks valid is that if Jim doubts B and believes A implies B then he certainly cannot believe A (supposing belief is closed under believed implications and Jim is rational). (One might recognize this principle as a thesis of Richard Kilvington.) But then he knows he must hold that A is false or dubitable, and if he holds that A is false—i.e. he *disbelieves* A—then, on certain accounts, it is dubitable. So in either case Jim ought to doubt A (or at least he may, though in this case the stronger obligation seems more appropriate). It looks then as if doubt, just like denial, is closed under (believed) *converse implication*: if B is dubitable and A implies B then A too is dubitable.

An account of doubt according to which disbelief implies dubitability is reasonable on some ordinary uses of the term, but not on all. There is a notion of doubt as *agnosticism* according to which disbelief *rules out* doubt, for if one disbelieves a statement she cannot also be agnostic with respect to it. (By being agnostic with respect to A I mean that neither A nor its negation is believed, i.e. that neither A is believed nor disbelieved.) I shall fix on this latter notion of doubt since it seems to be one that arises naturally in constructive mathematics or more generally theories in which assertion and denial are epistemically constrained.[10] A natural interpretation of the conditions under which a statement is assertible, deniable and dubitable—in the context of mathematics—is that the statement be proved, refuted or neither, respectively. That is, the following conditions plausibly constrain assertion, denial and doubt:

(Assert) A is assertible iff A is proved;

(Deny) A is deniable iff A is refuted;

(Doubt) A is dubitable iff A is undecided.

We need not think these constraints apply only to constructivists. Presumably they ought to apply to classicists as well. It would seem irrational for a classicist to assert (deny) Goldbach's conjecture without having actually obtained a proof (refutation). And without a proof that the conjecture is either provable or refutable one would not be warranted in asserting that that is assertible either. Of course that does not amount to denying an instance of excluded middle so the classicist is not in any trouble, and she may even still hold that the conjecture is either assertible or deniable if she holds that a sentence is assertible (deniable) iff it is true (false). But being assertible hardly seems equivalent to being true for truth seems *at most* a necessary condition on assertibility and not a sufficient one, and likewise for deniability and falsity.

Doubt as agnosticism is *not* closed under (believed) converse implication. Looking back at the example conversation we can see why. Jim might doubt

---

[10]In [Sal95], Salmon defines doubt in A disjunctively as either disbelief in or agnosticism with respect to A. This notion of doubt is weaker than doubt as agnosticism and, on the assumption that agents are consistent in their beliefs, equivalent to (modulo the usual normal modal logics of belief, e.g. **K4**) doubt as *non-belief*. For the most part I shall, but needn't, make this consistency assumption (typically encoded by the schema "D", $\Box A \rightarrow \Diamond A$), and so Salmon's disjunctive account turns out not very plausible.

that Sarah will be at the theatre while *disbelieving* that Paula will be. Since disbelief is incompatible with doubt in the present agnostic sense, Jim cannot doubt that Paula will be at the theatre. (Sometimes in contexts like the one Max and Jim are in, 'I doubt A' means 'I disbelieve A'. e.g. consider 'I doubt you can hit the bullseye!' uttered during a game of darts. In these cases doubt as disbelief *is* closed under converse implication.) So if the agnostic sense of doubt is neither constrained by implication nor converse implication, by which notion of consequence, if any, is it constrained?

In answering that question let us consider again the properties of being proved, refuted and undecided. Clearly they are pairwise mutually exclusive and jointly exhaustive over the set of all statements, i.e. every statement has precisely one of the properties. Suppose we take these properties as semantic values in a theory of assertion, denial and doubt. Perhaps surprisingly we find that they behave just like the values in a strong three-valued Kleene matrix. We have:

- $\neg A$ is undecided iff A is;

- $A \wedge B$ is undecided iff both A, B are, or if just one is then the other is not refuted;

- $A \vee B$ is undecided iff both A, B are, or if just one is then the other is not proved.

(Implication would be the only tricky case under a constructive notion of proof, but under a classical notion of proof we can simply define $A \rightarrow B$ in one of the usual ways, e.g. by $\neg(A \wedge \neg B)$, whereupon $A \rightarrow B$ would be undecided iff both A and $\neg B$ are, or if just one is then the other is not refuted. If one is inclined to think constructively about proof, then implication will be left out of the analogy between undefinedness in Kleene semantics and being undecided.)

Clearly the following constraints on assertibility and deniability hold.

(Assert-Cl) If A is assertible and it has been proved that A implies B then B is assertible.[11]

(Deny-Cl) If B is deniable and it has been proved that A implies B then A is deniable.

That is, assertion and denial are closed under proved implications and the converse of proved implications, respectively. Now what sort of analogous constraint could we place on dubitability? When the relevant notion of implication is the same for each assertibility, deniability and dubitability, then no such analogous constraint exists. We need a notion of implication as preservation of undecidedness.[12]

---

[11]It is important to note that 'assertibility' means '*may* be asserted' and not 'has been asserted' or 'must be asserted'. Obviously one may assert A without having any sort of obligation, even rational, to assert B when A implies B and even when he knows the implication holds, for in the former case maybe he has never even considered A, or in the latter case maybe it would take more than a lifetime to assert it.

[12]Suppose we had decided to opt for provability rather than provedness and likewise for refutedness and undecidedness. Then even for intuitionistic logic, the general case of condition (Deny-Cl) requires a consequence relation preserving refutability (in a multi-premise setting) that differs from the converse of the provability-preserving relation, since provability-preserving consequence does not anti-preserve refutability.

4.4 Doubt in Kleene logic

Let $\langle 0, 1, 2 \rangle$ be the order of truth values of the three-valued strong Kleene matrix K_3, where 0 is falsity, 1 undefinedness and 2 truth. Then let us call 2-consequence (or 2-implication) the relation of 2-preservation from sets of formulae to formulae defined over K_3, and 1-consequence and 0-consequence to be the relations of 1-preservation and 0-preservation similarly defined. I denote these relations respectively by $\vdash^2_{\mathbf{K_3}}$, $\vdash^1_{\mathbf{K_3}}$ and $\vdash^0_{\mathbf{K_3}}$. Since $\vdash^2_{\mathbf{K_3}}$ is well-known I refer to it also as simply **K3**. The use of the term 'consequence relation' to apply to these relations is justified since each of them satisfies overlap ($\Gamma \vdash A$ if $A \in \Gamma$), dilution ($\Gamma' \vdash A$ if $\Gamma \vdash A$ and $\Gamma \subseteq \Gamma'$), and cut (if $\Gamma \vdash A$ and $\Delta, A \vdash B$ then $\Gamma, \Delta \vdash B$). Sometimes such relations are called *Tarskian*. If we require also structurality for being Tarskian, then these relations are Tarskian in this stronger sense too. I am omitting set braces for singletons on the left of $\vdash$ as is standard, and the comma ',' is to be read as set-theoretic union on both the left and right of $\vdash$.

Given the above conditions (Assert), (Deny) and (Dobut), the following three constraints are each plausible:

(Assert-Cl) If A is assertible and A 2-implies B then B is assertible;

(Deny-Cl) If A is deniable and A 0-implies B then B is deniable;

(Doubt-Cl) If A is dubitable and A 1-implies B then B is dubitable.

There are a number of benefits of distinguishing a consequence relation $\vdash^s_{\mathbf{K_3}}$ for each truth value s. The first is that each constraint can be given in a uniform way: each of the conditions is subsumed under a general schema which we might call (X-Cl) for X a speech act or attitude. Second, if $\vdash$ is a truth-preserving consequence relation, it does not follow that its converse is a falsity-preserving consequence relation. In general it does not follow that it is an s-preserving consequence relation for any distinct semantic value s. This is true of $\vdash^2_{K_3}$ which anti-preserves *either* falsity *or* indeterminacy, but never always one or the other. Thus distinguishing various notions of consequence may in fact be a *necessity* when specifying constraints for certain speech acts or attitudes.

However, it will not always be a necessity. In certain cases, e.g. in classical logic, it is not a necessity since consequence anti-preserves falsity, but many logics are not like classical logic in this respect. For example, supervaluation-ist consequence fails to anti-preserve superfalsity (when consequence is taken "globally"), intuitionistic logic fails to anti-preserve refutability (in the general multi-premise case), the logic of paradox **LP** fails to anti-preserve falsity, and so on. This is why many non-classicists have been forced into distinguishing denial as a *sui generis* speech act that is not reducible to assertion of negation. They have had to give up what might be called 'Frege's thesis', viz. that denial just is the assertion of a negation, and they have had to give up the tight connection between negation, denial and falsity. For example, a gap theorist can deny A because it is gappy without asserting $\neg A$. Dually, a glut theorist can assert $\neg A$ because it is glutty without denying A.

So dubitability as undecidedness is constrained by 1-consequence. For a presentation of what 1-consequence looks like the reader is deferred to sec-

tion 4.6 where a natural deduction system is given and then proved strongly complete with respect to its intended semantics.

While Doubt-Cl provides a plausible constraint on doubt in mathematical discourse, on the assumption that doubt is in these circumstances epistemically constrained, does it also provide a plausible constraint for doubt as it occurs in other discourses where it may not be epistemically constrained? In general it does not, and the reason for this is tied to the reason "classical gap theorists", for whom doubt is not epistemically constrained, opt for supervaluationist semantics over strong Kleene semantics since only on the former do all classical principles come out valid.

On the last remark there is a caveat. Consequence may be defined in a number of ways. The standard is the truth-preservational definition according to which A is a consequence of Γ just in case every model in which each member of Γ is true, A is true too. Another is to define consequence as the impossibility of the truth of each of the premises together with the falsity of the conclusion. On this definition of consequence over K_3, all classical principles come out valid. I do not think consequence as truth-preservation has any priority over the "never true premises and false conclusion" definition. The fact that the two coincide for classical logic may be seen as showing that neither has priority. There is then an argument to be made for the strong Kleene semantics as a gappy semantics that a classical gap theorist could accept since it admits of formulae taking an intermediate truth value while holding that the tautologies are the formulae that are never false rather than those that are always true.

4.5 Doubt in modal logic

Atheists do not doubt that God exists, they *disbelieve* it. Suppose we wish to model this sense of 'doubt' so that 'I doubt God exists' comes out false when uttered by an atheist. A natural definition in a doxastic language with belief operators is given by

$$\triangle A := \Diamond A \wedge \Diamond \neg A.$$

It is easy to see that doubt, in this sense, is also not closed under converse implication. So, again, we need another notion of consequence under which doubt is closed.

Notice that $\triangle$ is just a contingency operator. Logics in which $\triangle$ is taken as primitive rather than definable in terms of a doxastic $\Box$ have been studied in the literature under the guise of logics for "contingency" or, equivalently, "non-contingency" where non-contingency ∇ is defined by $\nabla := \neg\triangle$. They were first investigated in [MR66] (though only for logics extending **KT**, i.e. non-doxastic logics), and later in [Mor76], [Cre88], [Hum95] and [Kuh95] (the latter two for logics not necessarily extending **KT**). Suppose our doxastic base for $\Box$ is **K4**. Then the logic **K4△** of doubt (or contingency, as it were) is the consequence relation $\vdash_{\mathbf{K4\triangle}}$ generated by the following axioms and rules.

- $\triangle A \equiv \triangle \neg A$

- $\nabla A \wedge \triangle(A \wedge B) \rightarrow \triangle B$

- $\nabla A \wedge \triangle(A \vee B) \rightarrow \nabla(\neg A \vee C)$

- If $\vdash A$ then $\vdash \nabla A$

- If $\vdash A \equiv B$ then $\vdash \nabla A \equiv \nabla B$

- If $\vdash A$ and $\vdash A \to B$ then $\vdash B$

K4△ is sound and complete with respect to the class of Kripke structures with a transitive accessibility relation.

The logic of $\triangle$ turns out interestingly similar to the logic of U in the signed proof system for $\vdash^1_{\mathbf{K_3}}$ (given below). For example the following characteristic principles hold:

- From $\triangle(A \wedge B)$ and $\triangle(A \vee B)$ it follows that $\triangle A$ and $\triangle B$;

- $\triangle A$ is interdeducible with $\triangle \neg A$;

- From $\triangle(A \otimes B)$ (for $\otimes \in \{\vee, \wedge\}$) either $\triangle A$ or $\triangle B$.

One crucial difference is that we no longer have that from $\triangle A$ and $\triangle B$ infer $\triangle(A \otimes B)$. A classicist (or supervaluationist) might rejoice—for just because both A and its negation are dubitable it does not follow that $A \vee \neg A$ is, as $\vdash^1_{\mathbf{K_3}}$ would have it.

So contingency logics can do double duty as logics of dubitability. Although this connection between contingency and dubitability has gone unnoticed in the literature, its application to the area of imposing rational constraints on propositional attitudes akin to doubting is certainly compelling. One should not be too surprised as a quick comparison of metaphysical and epistemic modalities reveals the same. Indeed, **S5** necessity has been doing double duty as an alethic and an epistemic operator for quite some time.

4.6 Consequence as 1-preservation

The following natural deduction system for Kleene's strong three-valued logic is a *signed* system in the sense that the formulae of the proof language are signed formulae of the object language, where intuitively each sign represents an assignment of truth value to the formulae to which it is appended. It is worth briefly remarking as to why a signed, rather than unsigned, natural deduction system was chosen. Typically the notions of consistency and satisfiability are equivalent, where a set is j-satisfiable if there is a model (valuation) in which each of its members is assigned the semantic value j, and a set is consistent if it does not contain some special subset of formulae (e.g. $\{\bot\}$ or the set of all formulae). In the case of 1-satisfiability there is no interesting notion of consistency, since *every* subset of formulae is 1-satisfiable (e.g. under a valuation which assigns to the atomic constituents of each member of the set the value 1). The problem of proving completeness for a logic in which every subset is consistent can be quite elusive.

Another nice feature of the signed system below is that, for the entire range of semantic values j, a consequence relation as j-preservation is easily determined. In the case of the Kleene matrix K_3, one may easily determine consequence as 0-, 1-, and 2-preservation defined over K_3. Such is done for 1-preservation in Definition 4.6.1 and can be easily adapted to j-preservation for each $0 \leq j \leq 2$.

We begin with some notational conventions. Signs F, U and T are also denoted by S_0, S_1 and S_2 respectively and i, j, and k denote integers in $\{0, 1, 2\}$.

We let A, B, etc. range over *unsigned* formulae of the object language, and ϕ, ψ, etc. range over *signed* formulae of the proof language. Similarly, Γ and Δ range over sets of unsigned formulae and Σ and Θ over sets of signed formulae. $S_j\Gamma$ is the result of replacing each $A \in \Gamma$ with S_jA, i.e. $S_j\Gamma = \{S_iA : A \in \Gamma\}$.

The system $\mathbf{K}^I$ has the following rules.

$$\frac{TA \qquad TB}{TA \wedge B}\ \wedge\mathrm{I} \qquad\qquad \frac{TA_1 \wedge A_2}{TA_i}\ \wedge\mathrm{E}$$

$$\frac{FA_i}{FA_1 \wedge A_2}\ F\wedge\mathrm{I} \qquad\qquad \frac{FA \wedge B \quad \overset{[FA]}{\overset{\vdots}{S_iC}} \quad \overset{[FB]}{\overset{\vdots}{S_iC}}}{S_iC}\ F\wedge\mathrm{E}$$

$$\frac{TA_i}{TA_1 \vee A_2}\ \vee\mathrm{I} \qquad\qquad \frac{TA \vee B \quad \overset{[TA]}{\overset{\vdots}{S_iC}} \quad \overset{[TB]}{\overset{\vdots}{S_iC}}}{S_iC}\ \vee\mathrm{E}$$

$$\frac{FA \qquad FB}{FA \vee B}\ F\vee\mathrm{I} \qquad\qquad \frac{FA_1 \vee A_2}{FA_i}\ F\vee\mathrm{E}$$

$$\frac{TA}{F\neg A}\ F\neg\mathrm{I} \qquad\qquad \frac{FA}{T\neg A}\ T\neg\mathrm{I}$$

$$\frac{S_i\neg\neg A}{S_iA}\ S_i\text{-DNE} \qquad\qquad \frac{\overset{[S_iA]}{\overset{\vdots}{\bowtie}} \quad \overset{[S_jA]}{\overset{\vdots}{\bowtie}}}{S_kA}\ \mathrm{RAA}$$

$$\frac{S_iA \qquad S_jA}{\bowtie}\ \bowtie\mathrm{I}$$

In RAA and $\bowtie$I the indices j, k and l are pairwise distinct and $\bowtie$ is an expression exclusively of the proof language which plays the role of absurdity. Thus it occurs only in applications of RAA and $\bowtie$I and nowhere else.

We define a relation $\vdash^U_{\mathbf{K_3}}$ (to correspond to $\models^1_{K_3}$) between sets of *unsigned* formulae and unsigned formulae in the following natural way:

Definition 4.6.1. $\Gamma \vdash^U_{K_3} A$ iff $U\Gamma \vdash_{K_3} UA$.

The relations $\vdash^T_{K_3}$ and $\vdash^F_{K_3}$ are defined in parallel fashion, substituting T and respectively F everywhere for U.

Completeness

Definition 4.6.2. Let Σ be a set of signed formulae. We say Σ is consistent if $\Sigma \nvdash \bowtie$ and Σ is maximal if for all formulae A, $S_jA \in \Sigma$ for some $j \leq 2$.

Lemma 4.6.3. *Every consistent set of signed formulae can be extended to (i.e. is included in) a Lindenbaum set (also called 'maximally consistent set').*

Proof. Suppose Σ is consistent and let $A_1, \ldots$ be an enumeration of all unsigned formulae. Let $\Sigma_0 = \Sigma$. Define recursively

$$\Sigma_n = \begin{cases} \Sigma_{n-1} \cup \{TA_n\} & \text{if it is consistent;} \\ \Sigma_{n-1} \cup \{FA_n\} & \text{if it is consistent;} \\ \Sigma_{n-1} \cup \{UA_n\} & \text{otherwise.} \end{cases}$$

Let $\Sigma' = \bigcup_{n \leq \omega} \Sigma_n$.

That Σ' is consistent follows immediately by construction. For maximality suppose for reductio that Σ' is not maximal. Then there is an A_n s.t. $S_j A_n \notin \Sigma'$ for each $j \leq 2$. By construction $\Sigma_{n-1} \cup \{A_n\}$ is inconsistent for each $j \leq 2$. By RAA, $\Sigma_{n-1} \vdash_{\mathbf{K_3}} S_j A_n$ for each $j \leq 2$. Whence by $\bowtie I$, Σ_{n-1} is inconsistent contra construction. $\square$

Lemma 4.6.4. *If $U\Gamma \nvdash_{K_3} UA$ then either $U\Gamma \cup \{TA\}$ or $U\Gamma \cup \{FA\}$ can be extended to a Lindenbaum set.*

Proof. Suppose $U\Gamma \nvdash_{K_3} UA$ and, for reductio, that neither $U\Gamma \cup \{TA\}$ nor $U\Gamma \cup \{FA\}$ is consistent. Then $U\Gamma, TA \vdash_{K_3} \bowtie$ and $U\Gamma, FA \vdash_{K_3} \bowtie$. By RAA, $U\Gamma \vdash_{K_3} UA$ contra hypothesis. Whence either $U\Gamma \cup \{TA\}$ or $I\Gamma \cup \{FA\}$ is consistent, so by Lemma 4.6.3 at least one of them can be extended to a Lindenbaum set. $\square$

Lemma 4.6.5 (Model existence)**.** *Let Σ be a Lindenbaum set. Then there is a K_3-valuation v s.t. for each A, $S_j A \in \Sigma$ iff $v(A) = j$ for $j \leq 2$.*

Proof. Suppose Σ is a Lindenbaum set. Define an assignment $\alpha : Prop \to \{0, 1, 2\}$ by

$$\alpha(p) = \begin{cases} 0 & \text{if } Fp \in \Sigma; \\ 1 & \text{if } Up \in \Sigma; \\ 2 & \text{otherwise, i.e. if } Tp \in \Sigma. \end{cases}$$

It is easily verified by the maximal consistency of Σ that α is well-defined (i.e. total and functional). We extend α to a unique valuation $v : Form \to \{0, 1, 2\}$ to arbitrary formulae according to usual recursive clauses for K_3. The rest of the proof proceeds by induction on formula complexity. The base case is trivial, so assume the inductive hypothesis (IH) holds for formulae of complexity less than A's. We do only the cases for $\neg$ and $\wedge$.

Case 1. $A := \neg B$. We have $T\neg B \in \Sigma$ iff $FB \in \Sigma$ iff, by IH, $v(B) = 0$ iff $v(\neg B) = 2$. Moreover, $F\neg B \in \Sigma$ iff $TB \in \Sigma$ iff, by IH, $v(B) = 2$ iff $v(\neg B) = 0$. Finally, $I\neg B \in \Sigma$ iff $IB \in \Sigma$ iff, by IH, $v(B) = 1$ iff $v(\neg B) = 1$.

Case 2. $A := B \wedge C$. We have $TB \wedge C \in \Sigma$ iff $TB, TC \in \Sigma$ iff, by IH, $v(B) = v(C) = 2$ iff $v(B \wedge C) = 2$. Moreover, $FB \wedge C \in \Sigma$ iff $FB \in \Sigma$ or $FC \in \Sigma$. W.l.o.g. suppose the former. Then $FB \in \Sigma$ iff, by IH, $v(B) = 1$ iff $V(B \wedge C) = 0$. Finally $IA \wedge B \in \Sigma$ iff either (i) $IA, IB \in \Sigma$, (ii) $IA, TB \in \Sigma$, (iii) $IB, TA \in \Sigma$. If (i) then by IH $v(A) = v(B) = 1$ iff $v(A \wedge B) = 1$. If (b) then by IH $v(A) = 1$ and $v(B) = 2$ iff $v(A \wedge B) = 1$. If (iii) then by IH $v(B) = 1$ and $v(C) = 2$ iff $v(A \wedge B) = 1$. $\square$

Theorem 4.6.6 (Completeness)**.** *If $\Gamma \models^1_{K_3} A$ then $\Gamma \vdash^U_{\mathbf{K_3}} A$.*

Proof. Suppose $\Gamma \not\vdash^U_{K_3} A$. Then either $U\Gamma \cup \{FA\}$ or $U\Gamma \cup \{TA\}$ can be extended to a Lindenbaum set Σ by Lemma 4.6.4 and by Definition 4.6.1. In either case it follows immediately by Lemma 4.6.5 that there is a valuation v s.t. $\forall \gamma \in \Gamma$, $v(\gamma) = 1$ and either $v(A) = 0$ or $v(A) = 2$. Whence $\Gamma \not\models^1_{K_3} A$. Contraposing, completeness follows. $\square$

4.7   Final remarks

We saw how to formulate rationality constraints on speech acts and attitudes, and in paritcular doubting, in a uniform way by distinguishing consequence $\vdash^s$ defined in terms of s-preservation for some given semantic value s. Thus if one has a particular attitude toward A and $A \vdash^s B$ then one ought to have the same attitude toward B. With regard to denial and rejection, s is often taken to be a given notion of falsity, e.g. superfalsity of supervaluationism, which may or may not be anti-preserved by consequence defined in terms of truth (e.g. supertruth) preservation as in the case of supervaluational consequence or consequence relations defined over an $n \geq 3$-valued semantics. What this highlights in the broad context of formulating rationality constraints on speech acts and propositional attitudes generally is the importance of, not just truth and truth-preserving consequence, but other truth values and consequence relations preserving just these values.

Chapter 5

On the nature of truth values

Abstract

Truth values are commonly employed in formal semantics for both logic and natural language, but if we suppose that correct semantical theories ought to be broadly representational then questions concerning the precise nature of truth values arise: Just what are they (representing)? I provide an answer to this question within a truthmaker framework. One surprising consequence of the account is that there is but a single truth value, truth. I finally consider some results of Suszko, Routley and Thomason adapted to the present case to show that every logic has a one-valued semantics, thereby vindicating the present philosophical defense of a single truth value.

5.1   Introduction

Truth values play a central role in contemporary semantical theories. If we suppose semantical theories ought to be broadly representational then questions concerning the precise nature of those values arise: just what are they (representing)? Frege took them to be the referents of sentences but besides their role in semantical theory, he said nothing (besides a vague remark) about their nature. In this paper I wish to elucidate their nature within the framework of truthmaker theory. One surprising consequence of the account is that there is but a single truth value, truth. If truth is the only truth value then false propositions do not refer to anything and it follows that a semantics, if it is to be genuinely representational, must be one-valued (or "monovalent"). This entails, again surprisingly, widespread semantic gappiness: i.e. the (intended) semantics for most logics, and this includes classical, intuitionistic and relevant logics, is gappy. In 5.4 I consider results of Suszko, Routley and Thomason adapted to the present case to show that every logic has a one-valued semantics, thereby vindicating the present philosophical defense of a single truth value.

The chapter proceeds as follows. In section 5.2 I discuss Frege's view of truth values as *sui generis* objects. I then give a philosophical defense of widespread semantic gappiness in section 5.3 backed up by the formal results of section 5.4. Finally I give an account of truth values in section 5.5 which rejects truthmaker maximalism (on which see section 5.6) in favor of a version restricted to *positive* propositions, a characterization of which is given in section 5.7. While negative truths fail to have truthmakers, the account of truthmaking given here is otherwise very inclusive, covering both universal and intensional positive propositions (sections 5.7 and 5.7) as having truthmakers of the most plausible sort.

5.2 Frege on truth values

Frege was the first to argue for truth values as *sui generis* objects rather than e.g. properties of sentences (or propositions and derivatively of sentences). He believed there to be precisely two of them, 'the true' and 'the false':

> By the truth value of a sentence I understand the circumstance that it is true or false. There are no further truth values. For brevity I call the one the true, the other the false. Every declarative sentence concerned with the referents of its words is therefore to be regarded as a proper name, and its referent, if it exists, is either the true or the false. [Fre48, p. 216]

The view that sentences *refer* to truth values follows naturally on Frege's view because he treated sentences on a par with proper names, each referring to saturated entities, unlike function expressions and (open) predicates which refer to unsaturated entities such as functions or concepts. Since proper names have both a sense and a referent, so too must sentences.[1] Now there are many candidates available to play the role of sentence referents, and Frege's choice of truth values or senses (the latter, e.g., in indirect speech reports) to play that role was based on two principal arguments. The first is that, he claims, "It is the striving for truth that drives us always to advance from the sense to the referent" [Fre48, p. 216].

The second argument he states thus:

> "The thought that 5 is a prime number is true" contains only a thought, and indeed the same thought as the simple "5 is a prime

---

[1]There is likely more to the story than this. For one, Frege held a principle of compositionality whereby sentences must have semantic values if they are to play any semantic contribution to the sentential wholes of which they form a part. Frege took these values to be truth values and senses. But the fact that the semantics requires sentences to have a semantic value does not lend compelling argument for the view that natural language sentences refer to truth values. In modern semantic theory for natural languages, the semantic value of quantifiers are, e.g., functions from properties to truth values or (in generalized quantifier theory) relations between relations defined over the domain. It does not follow, however, that we should take seriously the idea that quantifiers refer to anything let alone relations between relations defined over the actual domain. We might instead favor broadly truth conditional (e.g. Tarskian) semantics which endow sentences with meanings, not in virtue of assigning each of their subsentential parts semantic values as referents, but in virtue of their having truth conditions. The meaning of (non-sentential) subsentential expressions is the contribution they make to the truth conditions (or determination of such conditions) of sentences of which they form a part.

> number." It follows that the relation of the thought to the true may not be compared with that of subject to predicate. Subject and predicate (understood in the logical sense) are indeed elements of thought; they stand on the same level for knowledge. By combining subject and predicate, one reaches only a thought, never passes from sense to reference, never from a thought to its truth value. One moves at the same level but never advances from one level to the next. A truth value cannot be a part of a thought, any more than say the sun can, for it is not a sense but an object. [Fre48, p. 217]

These arguments are not persuasive to establish what Frege wanted, viz. that truth and falsity are objects rather than properties and, moreover, that sentences refer to them.[2] But I do think that Frege was right about the *kind* of thing truth values are—they are circumstances or states of affairs of a very particular sort, viz. *the* maximal sort (on which more later).

I wish to fill a major gap in Frege's argument by specifying what the circumstance that a sentence is true should mean, not according to Frege, but according to a truthmaker theorist. I follow Frege in holding that truth may be regarded as an object in its own right, but I am not wedded to the view that truth *must* be an object rather than a property (where properties are not themselves taken as objects in the relevant sense). If truth is an object—and this would require some argument—then the account given here provides a reduction of its status from primitive to something more intelligible, indeed something completely ordinary in kind.

If truth is an object and sentences are true in virtue of referring to or being assigned truth, then a truth property can be derivatively defined by saying that a proposition has the property just in case it refers to or is assigned the object truth. (Moving in the other direction, from the property to the object, is not always guaranteed; i.e. there may be a property of truth without any derivative sense of truth as an object.) So even taking truth primarily as an object, it still makes sense to speak of truth *derivatively* as a property of truthbearers. Truth-as-object talk does not rule out truth-as-property talk. Concerning whether sentences refer to anything, let alone truth values, will not concern me here.

---

[2]One obvious problem with Frege's argument for discounting truth as a property is that it begs the question; it presupposes that truth *is* an object referred to by sentences and that knowing that a sentence is true just comes to knowing the referent of the sentence. Another problem is that taking truth values as primitive does not tell us anything substantive about the nature of those values. While Frege does say that truth values are circumstances of a certain kind, this gloss is of little help in understanding what he meant. For the circumstance that a given truth (falsity) is true (false) is precisely the same circumstance that any other given truth (falsity) is true (false). A circumstance, in this sense, could only be something like a maximal state of affairs. But there is only *one* such state of affairs (assuming it is actual rather than merely possible) and there are two truth values, so a circumstance in Frege's sense cannot even be a maximal state of affairs unless there is, quite dubiously, something like a negative maximal state of affairs. Allowing for merely possible states of affairs will not help since there will be more than just two maximal ones—indeed there will be infinitely many—making any given sentence true or false. If one accepts Frege's distinction between concept and object (see [Fre92]) then the claim that sentences refer to truth values becomes more plausible.

5.3 Every logic is gappy

The notions of a gappy theory and a gappy logic (where a logic is understood as a multiple- or single-conclusion consequence relation and a theory is understood as a set of sentences closed under a given logic) vary greatly in the literature. On my favored conception of gappiness, *every* logic is gappy. But before moving to such a conception, first consider some of the usual frontrunners. (I will, for the moment, concentrate only on gappy *theories* rather than logics but a lot of what is said about theories can be said about logics as well, and in particular everything about semantic accounts of gappy theories discussed below.[3]) On one view a gappy theory contains, for some sentence A, neither A nor its negation $\neg A$—i.e. gappiness is incompleteness. But then incomplete theories containing excluded middle, viz. $A \vee \neg A$ for every A, turn out gappy and this seems counterintuitive especially if the intended semantics of the theory is bivalent. In the least, an account of gappiness should not *rule out* that theories for which excluded middle holds be gapless. Ideally it should be at least neutral with respect to such cases.

Another account takes a gappy theory to be one that does not contain $A \vee \neg A$ for some A. Accordingly, certain intuitionistic theories are gappy. But this is incompatible with plausible intuitionistic theories of truth. For let T be a truth predicate for the language of such a theory. Then granting plausible properties of the truth predicate, there is a sentence A such that $\neg T(\ulcorner A \urcorner) \wedge \neg T(\ulcorner \neg A \urcorner)$. By an intuitionistically acceptable De Morgan's law, $\neg(T(\ulcorner A \urcorner) \vee T(\ulcorner \neg A \urcorner))$, and by $T(\ulcorner \neg A \urcorner) \to \neg T(\ulcorner A \urcorner)$ and closure under substitution of provable implications we have $\neg(T(A) \vee \neg T(A))$, an intuitionistic contradiction. So intuitionistic logic is not gappy assuming some very plausible assumptions about truth.

Finally should we say a gappy theory is one couched in a language L whose semantics (intended or not) does not assign a semantic value (under some interpretation) to some sentence of L? Such an account renders any $n \geq 3$-valued semantics employing total valuation functions, such as the semantics for theories based on strong Kleene logic, gapless contrary to such theories being considered paradigm cases of gappy theories. An obvious modification to this account requires that the semantics be *intended*; i.e. a theory is gappy if its intended semantics (supposing it has one) has a value that is intended to *represent* a gap. But what does it mean for something to *represent* a gap if there is nothing to be represented? A piece of mathematics does representing according to whether there is something in the world it suitably matches with respect to certain relevant structural characteristics. A *truth value gap* does not exist so it does not have structural characteristics and things that do not exist cannot be represented. There just is no way to *represent*—and here I do not mean just *model*—a truth value gap.[4] The only plausible reworking of the last formulation of gappiness is to require that there be some *intended representational* semantics (I do not assume uniqueness of intended semantics, and it is important here that the semantics be both intended and representa-

---

[3]One may also take logics to be a species of theories if logics are defined as sets of sentences which are, by definition, closed under themselves. However, such a definition of a logic is not sufficiently discriminating.

[4]There is some debate about this when 'gap' means something like 'hole' and not 'truth value gap'. See e.g. [VC09].

tional) according to which some sentence *fails to receive a value* (under some interpretation). In the case of strong Kleene semantics, the total three-valued assignments ought to be replaced by partial two-valued assignments and the truth conditions of complex formulae modified in the obvious way. One then gets an equivalent semantics that is gappy in what I take to be the correct sense.[5]

(There are other accounts of gappiness that I will not discuss here, such as one which deems a theory gappy just in case it is not prime, i.e. in case it contains, for some A and some B, $A \vee B$ and yet it contains neither disjunct. Such an account is arguably much too inclusive. Another account, endorsed by Field [Fie08], deems a *truth* theory (e.g. **KFS**) gappy (distinct from his "paracomplete") just in case it contains $\neg(T(\ulcorner A \urcorner) \vee T(\ulcorner \neg A \urcorner))$ for some A. The account is plausible but limited since it applies only to truth theories couched in languages containing their own truth predicate. Also limited is an account that regards a theory as gappy if for some A neither A nor its (or a) negation $\neg A$ fail to receive a value for it requires of the language that it contain a suitable negation having certain properties. A semantic conception of gappiness is, thus, more flexible.)

But it is not just a reification of gaps that undermines the "semantics as representational" project. A reification of falsity is just as troubling, for there is nothing in reality which corresponds to falsity—not absences, lacks, negative facts or states of affairs or polarities. In particular, *classical logic is gappy*, for sentences that are not true should not be assigned any semantic value since falsity does not exist. Wittgenstein [Wit21, 4.43] made the observation that classical logic could make do with a one-valued semantics but it is not clear whether he drew any significant ontological consequences from the observation.[6] I shall draw such consequences here by arguing for an account of the nature of truth values from which it follows that only truth exists.

One might think the existence of a single truth value would seriously jeopardize the enterprise of working out representational semantics for non-classical logics that are typically characterized by semantics which posit more than one truth value. But, in the spirit of Suszko [Sus76] (or more accurately, a close cousin), we can think of the values assigned to sentences other than truth, not as truth values, but as mere "algebraic" or "semantic" values such as credences of belief. Whatever one thinks about Suszko's thesis restricted to a single logical value, every logic has a one-valued semantics.[7] This result follows from some trivial consequences of results found in [Sus76], [Tho76] and [Rou75], the philosophical upshots of which are defended in section 5.7.

---

[5]Kripke [Kri75, footnote 18] appears to take this view regarding his intended interpretation of the fixed-point construction, here being different from the semantics he actually employs, which he regards as not involving a truth value in addition to truth and falsity.

[6]Black [Bla64] calls the 'P-theory' one on which false propositions "picture" merely possible situations and true ones picture actual situations. If Wittgenstein held the P-theory then he could not have been considered an atomist in the sense most have in mind. On the other hand, if he rejected the P-theory, as Read [Rea05] argues, then he may be interpreted as having rejected a reification of falsity and having endorsed a gappy conception of classical truth-functional logic.

[7]As section 5.4 illustrates, all that is required of a relation to be a logic (i.e. consequence relation) is that it be reflexive (on which more in section 5.4).

5.4 One-valued semantics

Most semantics for formal languages are given in terms of an $n \geq 2$-valued semantics. It is not always clear that for such languages there exists an equivalent one-valued semantics. In general for any language given an n-valued semantics, it is not clear that there is an equivalent $m < n$-valued semantics, where equivalence means identity of consequence relations defined in the same manner, e.g. as preservation of a certain kind of semantic value, e.g. the designated ones. For instance, it is not clear that a finitely-valued semantics can be given for the usual continuum-valued Łukasiewicz logic.[8] It is clear that we can do so for classical logic by replacing the *total* assignment functions from atomic formulae to two truth values by *partial* assignments from atoms to a single truth value. We then get a gappy semantics for classical logic. To be specific, let $Prop = \{p_1, \dots\}$ be the set of atomic formulae of the language and let the class of assignments be the class of partial functions $f : Prop \to \{1\}$. Define recursively the class of valuations $v : Form \to \{1\}$ from arbitrary formulae to truth as follows:

- $v(A) = 1$ iff $f(A) = 1$ for $A \in Prop$

- $v(\neg A) = 1$ iff $v(A) \neq 1$

- $v(A \vee B) = 1$ iff $v(A) = 1$ or $v(B) = 1$

with the other connectives being defined in the usual way. (The metatheory is, importantly, classical rather than e.g. intuitionistic.) Presented with truth tables, we could simply replace all the 0s with blanks to show that the corresponding formula is not assigned a value, an idea that Wittgenstein had in mind in [Wit21, 4.43]. The same applies to a first-, or even higher-, order language given a truth-value semantics in the sense of [Leb76].

That the standard and "gappy" or partial semantics are equivalent is trivial. Obviously the same method can be extended to any Kripke semantics invoking two-valued valuations. In fact, when one takes valuations to be assignments of sets of worlds to atomic formulae, the semantics is effectively partial. (Truth at a world being membership in a set and falsity being lack of membership.) So one-valued semantics can be given for any Kripke logic characterized by a two-valued world semantics. The question is whether we can extend this method to $n \geq 2$-valued logics of *any* sort. For many-valued modal logics the result already follows trivially from [Tho76] (again by replacing the total two-valued functions by one-valued partial ones), but the following discussion highlights substantially more general results.

Suszko proposed the rather bold thesis that "Łukasiewicz is the chief perpetrator of a magnificent conceptual deceit lasting out in mathematical logic to the present day" [Sus76, p. 377]. The conceptual deceit was to pretend that there was anything more than two truth-values taken in an metaphysically or logically significant sense. Other values besides truth and falsity are what Suszko called mere "algebraic values".

---

[8]Semantics are, strictly speaking, for languages, not logics. Sometimes one speaks loosely as logics having semantics. The reason for this is that the semantics, with respect to which a notion of consequence is defined, may be (constructed particularly for the purposes of being) sound and complete for a given logic. In that sense we may speak of a semantics for a logic. Sometimes I speak in this way for the purposes of brevity.

Suszko's claim is supported by some very interesting formal results. He was the first to establish that every structural Tarskian consequence relation $\vdash$, i.e. any relation (between sets of formulae and formulae) closed under the following properties

Reflexivity. $\Gamma \vdash A$ if $A \in \Gamma$

Transitivity. $\Gamma \cup \{A\} \vdash B$ if $\Gamma \vdash A$ and $A \vdash B$

Monotonicity. $\Gamma \vdash A$ if $\Delta \vdash A$ and $\Delta \subseteq \Gamma$

Structurality. $\Gamma \vdash A$ only if $\Gamma' \vdash A'$ where Γ' and A' result from Γ and A respectively by replacing each occurrence of some atom in each member of Γ and in A by an arbitrary formula,

has a two-valued semantics. A proof of Suszko's theorem may be easily adapted to show that every structural Tarskian relation has a one-valued semantics.

Theorem 5.4.1. *Every structural Tarskian consequence relation has a one-valued semantics.*

Proof. Let $\mathbf{L}$ be a structural Tarskian relation in a language L. It was first proved by Ryszard Wójcicki (1970) that $\mathbf{L}$ is characterized by its Lindenbaum bundle $\Phi_{\mathbf{L}}$ defined by

$$\Phi_{\mathbf{L}} := \{(L, D = \{A \in L : \Delta \vdash A\}, O) : \Delta \subseteq L\}$$

where O is a set of operations and each structure in $\Phi_{\mathbf{L}}$ is similar to the language L.[9] For suppose $\Gamma \vdash A$ and $\forall B \in \Gamma, h(B) \in D$ for $h : L \to \phi$ a homomorphism with $\phi \in \Phi_{\mathbf{L}}$. Then $\forall B \in \Gamma, \Delta \vdash B$, so $\Delta \vdash \bigwedge \Gamma$. As $\bigwedge \Gamma \vdash A$, by cut we have $\Delta \vdash A$. So $h(A) \in D$ as desired.

Let C be any class of algebraic structures $\mathcal{M} = (M, D, f_1, \ldots, f_n)$ similar to L which characterizes $\vdash$, where M is the set of algebraic values and $D \subseteq A$ the set of designated values. (In particular C may be the Lindenbaum bundle of $\vdash$ which we know characterizes $\vdash$.) Let H be the class of homomorphisms $h : L \to M$ from L to (the carriers of) each $\mathcal{M}$. Define the class $V = \{v_h : L \to \{1\} : h \in H\}$ of one-valued valuations by

$$v_h(A) = \begin{cases} 1 & \text{if } h(A) \in D \\ \text{undefined} & \text{otherwise.} \end{cases}$$

We show that $\Gamma \vdash A$ iff $v_h(A) = 1$ whenever $v_h(B) = 1$ for each $B \in \Gamma$ and each $v_h \in V$.

Suppose $\Gamma \vdash A$ and that $v_h(B) = 1$ for each $B \in \Gamma$. Then $h(B) = 1$ for each $B \in \Gamma$. Since $\mathcal{M}$ characterizes $\vdash$, it follows that $h(A) = 1$. By definition, $v_h(A) = 1$ as desired.

Now suppose that for each $v_h \in V$ we have that $v_h(A) = 1$ whenever $v_h(B) = 1$ for each $B \in \Gamma$. Given the definition of V, $h(A) \in D$ whenever

---

[9]An algebraic structure is similar to a language if there is a bijection between the operations of the structure and the constants of the language which preserves arity. A class C of structures characterizes a consequence relation $\vdash$ (i.e. any relation between sets of formulae and formulae) if $\Gamma \vdash A$ iff $h(A) \in D$ whenever $h(B) \in D$ for each $B \in \Gamma$ and homomorphism h from the language of $\vdash$ to C.

$h(B) \in D$ for each $B \in \Gamma$. As $\mathcal{M}$ characterizes $\vdash$, we have that $\Gamma \vdash A$. This concludes the proof that every structural Tarskian relation is characterized by a one-valued semantics. $\square$

Suszko's result can be significantly strengthened. It was proved in [CCCM01] that the assumption of structurality is not needed, and in [Tsu98] that transitivity and monotonicity may be further dropped. Thus every consequence relation satisfying just reflexivity has (i.e. is characterized by) a two-valued semantics, so by replacing the total two-valued valuation in the proofs of [CCCM01] and [Tsu98] by partial one-valued ones, every such consequence relation has a one-valued semantics.

We can generalize even further. Routley [Rou75] showed that every λ-categorial language has a two-valued world semantics. By replacing in that proof the total valuations with partial one-valued ones it once again follows that every such language has a one-valued world semantics.

So we have seen that an exceptionally broad class of logics can be given one-valued semantics which, from a formal point of view, vindicates the Fundamental thesis (see section 5.6) that there is but a single truth value, truth. It is at least true from a formal point of view that only truth matters and the following sections aim to show the same is true from a philosophical point of view.

Truth is to be defined as the maximal truthmaker and negative propositions are taken care of by recursive truth conditions similar to an atomist account of truthmaking. The general strategy endorsed here for dealing with negative propositions is familiar by now but the details, to which we now turn, provide a plausible account of *why and how* negative truths should be dealt with in this manner, something most accounts which give up unrestricted maximalism fail to adequately address.

5.5   Truthmaking

The main idea behind truthmaking is that truth supervenes on, or is grounded in, reality. A proposition is true *because* something in the world necessitates its truth. (Some (e.g. [Rea10], [Dod02], and [Lew01]) think it is better to formulate truthmaking in terms of supervenience on *how* things exist rather than on *what* things exist. That seems plausible for a good number of propositions ordinarily considered, but it seems wrong for existential propositions (e.g. ⟨Hesperus exists⟩) and identities (e.g. ⟨Hesperus is Phosphorus⟩. Probably it is best to formulate truthmaking in terms of both supervenience on what and how things exist, but I shall ignore this issue here.) What the necessitation condition (called 'Necessitarianism' in [Arm04] and 'truthmaker essentialism' in [Par99]) amounts to may be easy or difficult to specify depending on one's view of (i) truthmaking and (ii) propositions. For example, if propositions make essential reference to worlds and times, then necessitation comes easy. For if the true proposition expressed by 'Fluffy is brown', relative to a given time t and world w is at least as discriminating as ⟨Fluffy is brown at t in w⟩ (where ⟨A⟩ refers to the proposition that A) then that particular slice of Fluffy, the t-slice-of-Fluffy-in-w, trivially necessitates the truth of the proposition. On the other hand, if propositions do not make explicit reference to worlds and times,

as Kaplan and others have forcefully argued, spelling out what necessitation amounts to becomes an important problem in truthmaker theory.[10]

For a taste, here are the most popular proposals on the table. One takes the notion of *x's necessitating the truth of a proposition* $\langle A \rangle$ as primitive, another as equivalent to $\Box(E!x \Rightarrow A)$ (possibly without the $\Box$ if $\Rightarrow$ already has strong enough modal force) where $\Rightarrow$ is an entailment operation, e.g. relevant entailment, and $E!x$ is a formalization of 'x exists', sometimes defined as $\exists y\, y = x$ (though this definition is inadequate in a number of formal semantics for modal languages), and another as being reducible or explainable in terms of other, arguably non-modal, notions such as intrinsic properties or (more arguably non-modal, but see [Fin94]) essences. Significant worries arise for each of these accounts. Ultimately I reject Necessitarianism for the reasons discussed in section 5.7.

A given truth may have many truthmakers, always a maximal one and sometimes a minimal one. A truthmaker α is *maximal for* $\langle A \rangle$ when it is a truthmaker for $\langle A \rangle$ and no proper part of α is also a truthmaker for $\langle A \rangle$, and it is *minimal for* $\langle A \rangle$ when it is a truthmaker for $\langle A \rangle$ and nothing of which it is a proper part is also a truthmaker for $\langle A \rangle$ (e.g. see [Arm04, pp. 19-20]). One equivalent way of formulating maximality is to say that α is a maximal truthmaker for $\langle A \rangle$ when it is one that makes true any proposition made true by any other truthmaker for $\langle A \rangle$. This definition does not rely on mereological properties of parthood. However, no dual formulation of minimality is equivalent with the one initially given except, perhaps, for atomic propositions.[11]

Some propositions have many minimal truthmakers and others may have none (under the assumption, e.g., that matter is indefinitely divisible). But every truth has a maximal truthmaker on the assumptions that (i) we have a suitably strong notion of mereological fusion at our disposal and (ii) Maximalism holds, i.e. that

Maximalism. For every true proposition $\langle A \rangle$ there is an object α that makes it true.

I reject Maximalism but accept a version restricted to positive truths.[12] We will

---

[10]One such argument, essentially due to Kaplan [Kap78], is that there would be no contingent propositions if propositions make essential reference to worlds and times. For it is true at some world that 'ϕ is true at w' iff (if and only if) it is true in every world, and likewise it is true at some time (and a given world) that 'ϕ is true at t' iff it is true at every time. Alethic and temporal operators would then have no effect on the truth value of propositions and every proposition would be necessarily and always true. Typically no precise formulation of necessitation is given and instead the more cautious tactic of taking 'necessitate' as primitive is taken (see e.g. [Arm04]).

[11]The dual formulation of minimality would be: α is a minimal truthmaker for $\langle A \rangle$ when it is a truthmaker for $\langle A \rangle$ and any other truthmaker for $\langle A \rangle$ is a truthmaker for any truth made true by α. To see that the two formulations of minimality are not equivalent consider the proposition that cats exist. Then Fluffy and Blackie are both truthmakers for that proposition, and on the original formulation, minimal truthmakers, though on the second formulation neither is a minimal truthmaker since Fluffy makes true $\langle$Fluffy exists$\rangle$ while Blackie does not.

[12]Lewis [Lew03, p. 29] says "[s]ome philosophers hold [Maximalism]: they say that every truth must have a truthmaker. That is, all propositions are positive." So he thinks Maximalism is equivalent to the claim that all propositions are positive. But Lewis does not think positivity and negativity are mutually exclusive, whereas I (and most) do. In fact, he thinks every proposition is both positive and negative. That is why Maximalism restricted to positive propositions does not amount to "Maximalism unrestricted" on my picture.

suppose for the moment that Maximalism holds since, for our purposes, what follows from Maximalism follows also from the version restricted to positive propositions.

Not only does it immediately follow that every truth has a unique maximal truthmaker, it follows that there is a unique maximal truthmaker that makes every truth true.

Proof. Consider all the true propositions. By Maximalism each such proposition, $\langle A \rangle$, has a truthmaker. Take all and only the truthmakers α_i (with $i \in I$ for some index I, e.g. I the class of all ordinals) for $\langle A \rangle$. Then the α_i *taken collectively* is a maximal truthmaker for $\langle A \rangle$, so there is a maximal truthmaker Φ_i for each truth $\langle A_i \rangle$. Take the fusion $\top$ of each Φ_i. Then $\top$ is a maximal truthmaker for each truth. For uniqueness, suppose β is also a maximal truthmaker for $\langle A \rangle$ that is distinct from $\top$. Then β is one of the α_i, so it is a proper part of $\top$. This contradicts the maximality of β, so $\top$ is unique. □

(The proof does not go through if 'taken collectively' does not mean 'fused' so that 'the α_i taken collectively' may refer to a plurality rather than a single individual. This would require giving up maximal truthmakers and revising Maximalism to state that for every true proposition there are *some objects*, rather than always a single one, that make it true. One might also reject the principle of *unrestricted* fusion—viz. that for any things x there is a (unique) fusion of those things—either because the principle is claimed to bear ontological weight, restricted or not, or because it ought to be restricted for other reasons, e.g. out of considerations of the individuation of ordinary objects. In regard to the former, it is not clear the principle does indeed bear ontological weight. Weak composition-as-identity theses, one version of which is defended in [Lew91], support the "ontological innocence" of the principle. If we assume unrestricted fusion, as I shall from hereon, Maximalism and its pluralized version (viz. for every true proposition $\langle A \rangle$, there *are some* objects α_i that make it true) turn out equivalent in which case the proof goes through regardless of whether Maximalism or the pluralized version is assumed.)

That there is a unique maximal truthmaker for every truth does not imply what Restall [Res96, p. 334] calls 'truthmaker monism', viz. that every truthmaker makes every truth true, for that requires further assumptions such as the classical entailment thesis. Nor does it trivialize the truthmaking enterprise. One of the main interests of truthmaking is to ground truth in *minimal*—or at least sufficiently small—truthmakers, the existence of which are not ruled out by the existence of maximal truthmakers. Similarly the fact that every explanation can be embedded in a stronger explanation (just conjoin a truth to the original explanation) does not trivialize explanations. Finally there is nothing like a slingshot argument available as an immediate consequence of UMT if the conclusion of that argument is that there is precisely one proposition. For propositions on the present account are structured entities, not truth values.[13]

On many accounts of propositions and truthmakers there is at least a proper class of propositions but far fewer truthmakers. If one is a nominalist about

---

[13]If the conclusion of the slingshot argument is that every truth corresponds to the same *fact* and 'fact' is understood as 'maximal sum' then the slingshot argument follows. But the slingshot argument is not typically understood this way because facts are not typically understood this way; they are not mere sums of objects, they are more like structured propositions.

abstract objects, then there are only as many truthmakers as there are concrete objects and compositions of those objects.[14] So we do not need to come up with an incredible amount of truthmakers for so many true propositions. Even fewer truthmakers may be posited if "the 'logical constants' are not representatives" ([Wit21, 4.0312]), i.e. if one holds an atomist theory of truthmaking. There is nothing problematic about providing truthmakers for certain compound propositions, such as conjunctions. The only problematic case involves negative propositions. One conclusion to draw from all of this is that the only constant that *must not* be a representative is negation and that one may, if one wishes, remain neutral as regards the others (assuming that from the others, taken together, negation is indefinable).

From what has been said about truthmaking so far, a proposition is true iff it has a truthmaker, iff it is made true by the maximal truthmaker. Thus being true and being made true by the maximal truthmaker are extensionally equivalent. If we restrict maximalism to positive propositions, there will be more to truth than being made true by the maximal truthmaker; in particular, negative truths will be true in virtue of lacking a truthmaker (or equivalently, lacking the maximal truthmaker). On the other hand suppose $\langle\neg A\rangle$ is a negative truth. Then $\langle A\rangle$ is a falsity on a standard-going definition of falsity. So if we discard all talk of negative truths in favor of talk of falsities then it remains true that all there is to truth is being made true by the maximal truthmaker (and all there is to falsity is lacking the maximal truthmaker).

Now if there is nothing more to truth than being made true by the maximal truthmaker, it follows that

Fundamental thesis. There is but one truth value, truth, i.e. the maximal truthmaker.

Truth is what Frege called the circumstance that a truth, any truth, is true. Having sympathies with Frege, we may think of 'α makes true $\langle A\rangle$' meaning nothing more than 'α is the circumstance that $\langle A\rangle$ is true' or even (and less circularly, since no reference to truth is made) 'α is referred to by A'.

Is there anything analogous to the circumstance that a falsity is false, such as the maximal falsitymaker? It is often supposed that α is a falsitymaker for a proposition iff it is a truthmaker for its contradictory. On the present view this must obviously be rejected since then the maximal falsitymaker is identical to the maximal truthmaker, in which case the absurdity that a sentence is true iff it is false follows. But that is what we would expect on the present account: there are no such things as falsitymakers and a negative proposition $\langle\neg A\rangle$ is not made true by $\langle A\rangle$ having a falsitymaker, it is made true wholly in virtue of lacking a truthmaker.[15]

---

[14]Whether or not the fusion of some objects is to be counted as an additional object is often thought to be a matter of whether certain forms of composition-as-identity hold. However, Baxter [Bax88] provides an argument against fusions counting as additional objects over their parts without even invoking composition-as-identity.

[15]Another argument against the existence of falsitymakers is that it assumes that the existence of certain objects x rules out by metaphysical necessity the existence of certain other objects y whose existence seem otherwise independent of the existence of the xs. For instance consider the truthmaker Fluffy for $\langle$Fluffy exists$\rangle$. Now it is not clear what would count as a falsitymaker for the proposition (assuming that such things existed), but whatever those things are (e.g. a totality fact fused with some other object, a negative fact, etc.), their existence is, at least *prima facie* implausibly, ruled out by the existence of Fluffy.

I have sketched a view of the ontology of truth values which requires a rejection of Maximalism and an account of positive propositions. I take up each task in turn.

5.6 Maximalism

Truthmaker theory is typically restricted to contingent, or at least non-logically true, propositions.[16] The reasons for doing so are twofold: either necessary truths are made true by anything and everything or they are made true by nothing. In either case truthmaking for necessary truths becomes a trivial matter, so the attention has focused on contingent truths.

Why should we think necessary truths are made true by everything? For one, it follows straightforwardly from a naïve view of truthmaking according to which α is a truthmaker for $\langle A \rangle$ iff the proposition that α exists implies $\langle A \rangle$, where implication need not be classical. (If one does not like to talk about implication as a relation between propositions, then rephrase everything just said in terms of *suitable* sentences expressing those propositions.) It also follows straightforwardly from Maximalism, even when restricted to contingent propositions, and the entailment thesis formulated using certain (e.g. classical or strict), but not all (e.g. relevant), implications. So there are grounds for holding that necessary truths are made true by everything. However such a move can also be easily resisted just by taking the truthmaking relation as primitive and denying that it is *in virtue of the nature* of any particular object that a given necessary truth is true. For a popular view is that one of the distinguishing features of necessary, especially logical, truths is that they seem not to be grounded in the way any particular world is, i.e. their truth is independent of "reality". This view, though popular, seems obviously fallacious when truthbearers are taken to be propositions, and I suspect it has become mainstream because what are usually taken to be necessary are not propositions but rather (interpreted) sentences or the like.

For suppose we grant there are *a posteriori* necessities such as $\langle$Hesperus is Phosphorus$\rangle$.[17] Since its truth depends on Venus—for if Venus did not exist neither would the proposition—it seems it has as a (minimal) truthmaker Venus. The only reason to think that it lacks a truthmaker is to think that it has the form of the logical (by which I mean *schematic*) truth $x = x$, as it does on a structured account of propositions, and that logical truths lack truthmakers. But we just noted that the truth of the proposition is grounded

---

[16]For example, see [Lew01], [Smi99], [Mol00], [Dod07] and (less obviously) [MSS84] and [Tal09].

[17]I am supposing the proposition is true even in worlds in which Venus fails to exist (assuming there are such worlds), otherwise it seems unlikely that there would be any *a posteriori* necessities, for they would have to make essential reference to dubious necessary existents. For if a contingently exists then $\phi(a)$ cannot be true at any world at which a fails to exist, for it would not express any proposition. (Of course one can deny this by denying the semantic assumption that sentences containing irreferential singular terms are never true, but this seems highly implausible on a structured view of propositions unless one holds that there are such things as gappy propositions (see e.g. [Bra93]) or that propositions, construed or represented as set-theoretic objects, may exist in worlds in which their members fail to exist.) Most grant *a posteriori* necessities and, in any case, broadening our use of 'necessity' to mean 'true in all worlds where the individuals referred to by the proposition exist' would make precisely the same point, so my assumption serves only to simplify matters.

in that part of reality that includes Venus regardless of its being a logical truth, for if Venus had not existed the proposition would not have existed and hence could not have been true, let alone necessarily true. So at least some *a posteriori* necessities have truthmakers.

Suppose we reject the above argument and maintain that, while the existence of ⟨Hesperus is Phosphorus⟩ depends on Venus, its truth does not. Does it follow, then, that it lacks a truthmaker? In general, what reason is there for thinking that logical truths lack truthmakers? One is that such truths are knowable *a priori*. But that is not good enough reason if there are *a priori* contingencies, for certainly those will have truthmakers. And even if there are no *a priori* contingencies, there still is not any relevant connection between a proposition having the form of a logical truth and it being knowable *a priori*.

For one, it is dubious that what are knowable *a priori* are things like structured propositions. There is compelling reason for the view that the objects of propositional attitudes like knowledge are, not entities like propositions, but rather hyperintensional entities like structured propositions plus modes of presentation δ, such as ⟨Venus, Venus, identity, δ⟩. Sentential contexts involving propositional attitudes are hyperintensional in the sense that substitution into them of necessarily equivalent propositions *salva veritate* fails whereas propositions and facts seem not to be like this, i.e. substitution *salva veritate* of necessary equals into the context 'It is a fact that...' holds (see Davidson 1969). This implies that the objects of propositional attitudes cannot be propositions. So while a proposition such as ⟨Hesperus is Phosphorus⟩ might have the form of a logical truth, it does not follow that any hyperintensional object corresponding to the proposition (such as ⟨Venus, Venus, identity, 'Hesperus = Phosphorus'⟩) will be knowable *a priori*—that will rest crucially on what the mode of presentation is. The fact that a proposition has the form of a logical truth does not by itself make it any more plausible that it should lack a truthmaker simply on the grounds that logical truths are knowable *a priori*.

Even though the above considerations regarding necessary truths carry over to logical and analytic truths as well, there are also independent reasons for holding that the latter have truthmakers just as much as the contingent truths do. Consider the analytic truth ⟨Bachelors are unmarried⟩. The proposition connects or unifies (unification being a *metaphysical* relation) bachelorhood and the property of being unmarried. The truth of the proposition is independent of the meaning of the words 'bachelors', 'are' and 'ummarried'. Now it is true that the proposition has a special *semantic* status—some relation of "containment" holds between the meanings or concepts it involves. But having such a status tells us nothing about the *non-semantic* fact concerning whether the proposition has a truthmaker. So to exclude analytic truths as having truthmakers strictly on the grounds of their semantic status is greatly mistaken.

For Maximalism

In this section I wish to look at what I take to be the most persuasive arguments for Maximalism and show them to be wanting. If I am right, Maximalism is not well-motivated. But then I see it as no mark against a truthmaker theory for rejecting Maximalism if the rejection falls out of the account as a natural consequence. Of course the theory might be criticizable on other grounds.

Armstrong [Arm04] argues for Maximalism by arguing that truthmakers can be found for all kinds of truths, such as necessary and negative ones.[18] But there are a good number of reasons—which I will not rehearse here but see e.g. [Mol00]—for not liking his proposal, especially his "totality facts" solution to the worry about truthmakers for negative truths. And if his proposal fails to give plausible truthmakers for *all* truths, as I think it does, then his argument for Maximalism collapses with it.

Cameron [Cam08] gives the following argument for Maximalism:

> The thought that we do not need truthmakers for negative truths does not seem much of a runner. But if that is right then the thought that we cannot have truthmakers for negative truths looks very serious indeed; for if true it would seem to motivate abandoning not just truthmaker maximalism, but truthmaker theory altogether. If we do not get the negative truths for free given the positive truths, then what possible motivation could there be for accepting that some truths require truthmakers but that negative truths do not? That would be to accept that the negative truths are not true in virtue of anything: but if we allow that then why do we not allow positive truths that are not true in virtue of anything? It is one thing to say that certain truths are obtained for free given our grounding of other truths, and hence that they do not need a further grounding; it is another thing altogether to say that certain truths just are not grounded. Either there is something wrong with accepting truths that do not have an ontological grounding or there is not: if there is, then every truth requires a grounding; if there is not, then no truth. [Cam08, pp. 411–412]

Cameron's argument brings to bear an important distinction between the relation *x is true in virtue of y* and the relation *x is ontologically grounded in y*, a distinction important for atomists and others (e.g. [Mel03] and [Mum05]) who deny that, for example, negative truths are ontologically grounded and yet accept that they are nonetheless true in virtue of *something* where that something may be semantic or logical facts.[19] Being true in virtue of something does not imply being ontologically grounded in something. Even if one denies any such distinction between the two relations, Cameron's argument (which might be seen as fallaciously inferring a $\forall\neg A$ from a $\neg\forall A$, though matters are not so simple here) would serve as an argument against all kinds of widely held views

---

[18]He also makes the claim that "One can, of course, simply assert that a proposition such as ⟨There are no unicorns⟩ stands in no need of any truthmaker or other ontological ground. But this seems to be no more than giving up on truthmakers as soon as the going gets hard" [Arm04, p. 70]. But the aim was not to motivate Maximalism on grounds *independent* of whether *plausible* candidates for truthmakers can be found. The original aim was to motivate Maximalism by finding said candidates and it is thus very much dependent on whether plausible candidates may be found for all species of truth. If the task of finding truthmakers for some species of truth (e.g. negative truths) turns out looking hopeless then that is reason to give up Maximalism given the Armstrong's original motivation. Current accounts of truthmakers for negative truths are not compelling enough to make a strong enough case for Armstrong's defense of Maximalism.

[19]For atomists, the complex truths are not grounded in the world, rather they are true in virtue of the truth value of the atomic propositions which occur in them (and in particular the grounding of those that are true) plus the churning out of some recursive semantic machinery on those atomic propositions. See e.g. [MSS84], [Smi99] and [Sim05].

across philosophy; in particular any view which endorses a restricted principle. For instance, it would serve as an argument against non-naïve set theories (if only some collections are not legitimate then none are), truth theories with a restricted T-schema (if the T-schema does not hold for all sentences then it holds for none), and so on. So the argument is obviously far too powerful. It may not always be *ad hoc* or unprincipled to restrict a given principle such as Maximalism. Indeed there are two very compelling reasons for restricting it: (i) there are no plausible truthmaker candidates for negative truths and (ii) there is no independent, compelling motivation to accept Maximalism (on which more in section 5.6). A thorough defense of (i) may be found in [Mol00].

Dodd [Dod07] claims that a rejection of Maximalism "undercuts the motivation for truthmaker theory in the first place" and provides only the following argument for it:

> The intuition that truth must be ontologically grounded in the sense explicated by [Maximalism] is an intuition concerning (non-analytic) truth in general: it is one particular way of trying to explain the intuition that what is true is determined by how things are, but not vice versa. Consequently, if it really is the case that this asymmetry can only be adequately explained by adopting a truthmaker principle, it would seem to be a failure of nerve to depart from this general principle in the wake of the problem of finding truthmakers for negative truths" [Dod07, p. 394]

The problem with this argument is that the asymmetry is not *explained* by Maximalism—it is explained by the truthmaking relation which is itself elucidated quite independently of whether or not Maximalism holds. The reason we think the world and its parts make propositions true and not conversely is that a notion akin to ontological dependence is built into the notion of truthmaking. Indeed, (if one holds a structured theory of propositions then) propositions themselves depend ontologically on the world since propositions are comprised of worldly parts and their truth depends on how those parts are unified to form facts.

Against Maximalism

I briefly mentioned theories of truthmaking based on logical atomism. We may also list Mellor's [Mel03] among such theories. He says

> Even some contingent truths need no truthmakers, notably true truth-functions, whose truth follows from the truth values of their constituents. We may *say* of course that 'P & Q' and '$P \vee Q$' are 'made true' by the truth of 'P' and 'Q'; but this is just the entailment of one proposition by others, not the "cross-categorical" link between propositions and other entities that concerns us here. That is what true truth-functions do not need and therefore, I claim, do not have. [Mel03, p. 213]

Other theorists reject Maximalism without necessarily endorsing atomism. For instance, Mumford [Mum05] denies that negative truths have truthmakers using a deflationary argument:

> I am tempted to treat all putative cases of negative truths as cases
> of falsehood, for example, to understand ⟨the door is not blue⟩ as:
> ⟨the door is blue⟩ is false. We can then say that ⟨q⟩ is false, $f\langle q \rangle$,
> means that there is no truthmaker for ⟨q⟩. This seems to me to be
> the simplest and most intuitive account of falsehood. [Mum05, p.
> 266]

A reinterpretation of Mumford's view that does away with the ontologically
loaded notion of falsity replaces that notion with *deniability*. For example,
we might be tempted to understand an utterance of 'The door is not blue' as
expressing a denial of the proposition ⟨The door is blue⟩ rather than expressing
an assertion of ⟨The door is not blue⟩. Then the denial is correct just in case
the proposition denied has no truthmaker.

Maximalism may be rejected for a number of good reasons, but I think the
most compelling reason is that there simply are no plausible candidates to play
the role of truthmakers for negative truths. Moreover, it just strikes me as more
natural to count negative truths true in virtue of lacking truthmakers than to
count them true in virtue of being made true by some ontological oddity such
as a totality or negative fact.[20] Indeed, I should hope it strikes even those
who advocate Maximalism as more natural despite their being theoretically
driven to the opposite conclusion on what I take to be uncompelling grounds,
but ultimately arguments from naturalness are not going to do a whole lot of
convincing in any case.

In the next section I give an account of positive propositions which puts the
necessary flesh on the defense of Maximalism restricted to positive propositions.

5.7 Positive propositions

Many have expressed doubt in the distinction between positive and negative
propositions. For instance, Anderson and Belnap [AB63, p. 304] claim that
"we have no *semantic* grounds for distinguishing between "positive" and "neg-
ative" propositions" and Ayer [Aye52] considers a number of ways of making
the distinction each of which appears to be faced with insuperable difficulties.
Certainly if what is negative is a matter of *syntax* then there is no problem
in distinguishing the positive from the negative. But propositions are not sen-
tences and whether a proposition, not a sentence, is positive or negative is not
a matter of syntax. Yet if we believe that there is a special class of proper-
ties which carve nature at its joints, call these properties 'fundamental', and
that a complete description of the world would involve only *simple* proposi-
tions of this sort, i.e. propositions not involving anything corresponding to
propositional operations such as implication, there appears to be an intuitive
account of the distinction between negative and positive propositions we seek.

---

[20]Jago [Jag10] takes a positive fact to be a non-mereological composition of a thin par-
ticular and a universal. He then says that a "negative fact would be what obtains when the
relevant particulars lack the relevant universal. So we can say, in parallel to what we said
about positive facts: what a negative fact is, is simply the relevant thin particular(s) and
universal, composed in a certain (non-mereological) way" [Jag10, p. 3]. The problem is that
types of facts are traded in for types of composition, the latter being taken as primitive and
the problem of explaining what a negative fact is gets pushed "down" to explaining what the
various modes of composition are. As they are taken as primitive, Jago has nothing to say
about this.

A proposition is positive if it attributes a fundamental property to an individual and it is negative if it denies a fundamental property of an individual. As it stands the account is incomplete—what do we say about propositions that do neither?—but as a first pass it is on the right track.

Indeed, Ayer endorses a story along these lines:

> . . . we are surely entitled to say that a statement is affirmative if it refers to some particular individual and ascribes a property to it, or if it refers to a particular set of individuals and states that they stand in a certain relationship, or if it asserts that some unidentified individual, or set of individuals, answers to a certain description. No doubt there are more complicated possibilities, the existence of which would make it a laborious matter to give a complete list of all the forms of affirmative statement. [Aye52, p. 802]

In what follows I undertake the "laborious" task of offering a completely general account of the distinction between positive and negative propositions which encompasses not just truth-functional compounds but quantified and intensional propositions as well.

Atomicity

There are two kinds of properties in the world: the fundamental ones and the abundant ones—at least, let us assume the distinction holds good. If there are not many simples in the world, e.g. if there are just space-time points, then there will not be many fundamental properties. Generally an inflation of simples yields an inflation of fundamental properties. It does not matter to me how one analyzes fundamentality if he analyzes it all. One may wish, as I prefer, to take the fundamental-nonfundamental distinction as primitive.[21] However one wants to make the distinction, it is the fundamental properties that underly the account of positivity to be given in what follows.

Molnar [Mol00] claims that "everything that exists is positive". That depends on what is meant by 'exist' and by 'positive'. We can be quite relaxed about what we take to exist, for instance, we might think it acceptable to include properties taken in the abundant sense. After all, I see no reason to think properties do not come cheap and that, at the very least, every well-formed predicate in our language corresponds to a property.[22] If we take this relaxed

---

[21]Sider [Sid96] has objected to taking the distinction as primitive on a class nominalist view of relations, but that objection relies on identifying properties with sets of ordered pairs, the latter themselves being *identified* with sets. I make no such identification even if I think ordered pairs may be adequately *represented* by sets only when we are concerned with a particular class of sentences about or properties of pairs. Such a class would *not* include e.g. 'Is the empty set a member of the pair (x, y)?' which would be sensitive to the set-theoretic way of representing tuples. Any question which is sensitive to *artifacts* of the representation would be barred from the class.

[22]One may immediately shutter at the thought of paradoxes. Is there really the property of being a non-self-membered set? I do not see why not. The problem is that forming the collection of all such sets does not yield a set—it is too big. In any case, if one does not like the limitation-of-size solution to the set-theoretic paradoxes she can choose her favorite. If she has no favorite then she can restrict my claim trivially to exclude problematic cases by requiring that properties only exist when their corresponding predicates are not only well-formed but also consistently axiomatizable relative to a given (accepted) background theory. Even with such a restriction in place, abundant properties will still come cheaply.

view of existence and of properties then Molnar's claim is too strong, for there are obviously negative properties that exist—e.g. the property of being distinct from oneself—so not everything that exists is positive. (At least distinctness seems a negative property since, apparently, the only way to define it requires the use of operators from which negation can be defined.) Molnar's claim that everything that exists is positive is incompatible with most theories of properties which allow for properties of the abundant type, a class of theories that should not be ruled out from one's account of existence or positivity. As such, an account of positivity, in the very least, ought to remain neutral on whether everything is positive.

However, we can weaken Molnar's claim to the claim that everything that is *simple or fundamental* is positive, and this will be true even with abundant properties around. Moreover, we can simplify a little by saying that an individual, rather than property, is positive only if it it exemplifies a fundamental property, since on many accounts composites do not exemplify fundamental properties. Thus

Simple non-propositional positivity. Something is positive *only if* it is a fundamental property or exemplifies a fundamental property.

The formulation of positiveness just given is precise but lacks detail. A filling in of the details requires giving an account of what constitutes simples and their fundamental properties. I will not be doing any such filling in here so that one is free to choose her theory of simples and fundamental properties as she wishes.

Now that we have a precise formulation of positivity (for individuals and properties), it is not difficult to give one for atomicity:

Atomicity. A proposition is atomic iff it (is represented by) a tuple whose form is $(\alpha_1, \ldots, \alpha_n, P)$, where each of its elements are positive or, in other words, the α_i are simples and P is an n-place fundamental property.

(I am not using 'property' to mean 'monadic property'—properties may be of any arity. Moreover, I am assuming a structured view of propositions as structured entities to be *represented by*—but not *identical to*—set-theoretic objects. Whether or not anything hinges on this depends on whether a plausible account of positivity can be given to unstructured propositions, a question I leave open for another occasion.) We may extend the notion of atomicity to cover composites and their properties, e.g. by allowing the α_i to be composites and P to be "nearly fundamental" assuming fundamentality admits of degree (see e.g. [Arm89] and [Lew86]), but for present purposes I prefer to relax the notions of a simple and a fundamental property. For instance, we may wish to grant physical atoms as being simple and the property of having atomic mass n as being fundamental and we may even get carried away and allow things and properties of the garden variety to count as simples and fundamental if we wished to extend atomicity to propositions including things of the garden variety. Such an extension would count e.g. ⟨This table is red⟩ as atomic without having to complicate the notion of atomicity by extending it beyond genuine simples. Such complications would only distract from the present point here so the notions of simplicity and fundamentality should be understood in a relaxed sense.

On the most austere conception of simples and fundamentals, not many propositions count as atomic on this account. On the usual atomist picture there is a close relationship between atomic propositions and atomic sentences as they are typically construed in formal languages. But I doubt any such relationship exists unless the structure of propositions very closely matches the structure of (formally regimented) sentences which represent them.[23] This seems unlikely since there are more propositions then there are sentences, in which case one wonders what the structure of these unexpressed propositions might be. Moreover it seems unlikely, from a metaphysical point of view, that the atomicity of *propositions*, taken as consisting of parts of reality, should depend essentially on syntax, something which does not appear to determine how or what things exist (besides, of course, syntactical things).

The account of atomic propositions just sketched is not strictly Wittgensteinian. For it is compatible with there being an inconsistent set of atomic propositions and logical properties, such as entailment, holding between atomic propositions. This is strictly forbidden on Wittgenstein's picture ([Wit21, 4.211, 5.134]), a picture I do not find compelling. Allowing sets of inconsistent atomic propositions does not imply that at least one of the members of the set must be negative. For example, nothing suggests that one of ⟨The flower is red⟩ and ⟨The flower is blue⟩ is negative simply because the pair is mutually inconsistent (intuitively or relative to a background theory of color). In other words, a set of atomic propositions may be inconsistent even though each member of the set is positive.[24]

Quantified propositions and necessitarianism

Atomic propositions are positive since they are composed of only positive things. A broader notion of positivity, which covers not just truth-functional compounds but compounds more generally, can be given by defining operations on properties and propositions which form positive properties and propositions. For example, conjunction forms positive propositions only from positive conjuncts, disjunction forms negative propositions only when at least one disjunct is negative, and negation forms positive propositions only from negative ones. This covers all the truth-functional compounds.

Now if finitary disjunction and conjunction form positive propositions from positive ones and negative ones from negative ones then there is no reason to think their infinitary generalizations do not do the same. If (following e.g. [Wit21]) a universal proposition just is (up to logical equivalence) the possibly

---

[23]Such a view of structured propositions is defended in [Kin95]. One immediate and forceful objection to the view is that it gives an account of propositions that is too fine-grained and this raises, among others, issues concerning shared content and disagreement. For instance the propositions expressed by 'John likes Sarah' and 'Sarah is liked by John' would be distinct.

[24]This should not be surprising. In most logics, not *theories*, any set of atomic propositions is consistent. However, this is not true for theories. e.g. in Peano arithmetic, the set $\{0 < 1, 1 < 0\}$ is inconsistent. (Even singleton sets are, e.g. $\{0 = 1\}$.) The reason this does not pose a problem for Wittgenstein when he remarks that any set of atoms is consistent is that he took atoms to "assert the existence of states of affairs" while he took mathematical and logical statements to be senseless since they failed to do any representing. Dummett, contrary to Wittgenstein, thought the set of all atoms (in a suitably interpreted language) to be inconsistent but for reasons other than those presently considered.

infinitary conjunction $\langle \bigwedge_{i \in I} P(a_i) \rangle$, where the a_i exhaust the domain of individuals at a given world, and an existential proposition just is (up to logical equivalence) the possibly infinitary disjunction $\langle \bigvee_{i \in I} P(a_i) \rangle$, then we have an account of positivity for universal and existential propositions.

Universal positivity. A universal proposition $\langle \forall x A(x) \rangle$ is positive just in case each of its instances $\langle P(a) \rangle$ is positive.

Call this the 'boolean account of quantification' or just 'the boolean account' for short.

There is a usual worry with this account. Suppose universally quantified propositions just are infinitary conjunctions and that α is a truthmaker for $\langle \forall x A \rangle$. One might ask in what sense α *necessitates* the truth of $\langle \forall x A \rangle$? It seems possible that the truthmaker for each conjunct necessitates that conjunct, and hence that α (e.g. the truthmaker for each conjunct taken collectively) necessitates the infinitary conjunction, without α necessitating the universal proposition. Supposing that A is not necessary, we could add something else to the domain whose being such that A is not necessitated by α, the reasoning goes. But then α does not necessitate the truth of the universal and hence the truthmaker for the infinitary conjunction need not be a truthmaker for the universal.

One might attempt to save the boolean account thusly. An infinitary conjunction is equivalent to the negation of an infinitary disjunction, and the negation of a disjunction (infinitary or not) is true just in case the disjunction lacks a truthmaker, i.e. just in case none of its disjuncts has a truthmaker. So a universal proposition is true just in case a certain proposition lacks a truthmaker. But this only holds if the disjunction is positive, otherwise we could rerun the same reasoning for an arbitrary proposition since (classically) any proposition is equivalent to a negation and we would not need to find truthmakers for any proposition at all. Though every proposition is equivalent to a negation, not every proposition is equivalent to a *negative* one (just consider the double negation of a positive proposition which is positive). For example, consider $\langle$Everything is spatially extended$\rangle$. It is positive if each of its instances is (and we are assuming they are). But then the universal proposition, construed as the negation of an infinitary disjunction, will be the negation of a negative proposition, viz. $\langle \neg \bigwedge S(a) \rangle$ (where $S(x)$ expresses that x is spatially extended), so it will be a positive proposition and a truthmaker for it must then exist.

I am inclined to just give up on necessitation. When we ordinarily speak of things making propositions true it is rarely the case that this holds of necessity *in the broadest sense*. Consider glass and its fragility. What makes it fragile? Well that has to do with its physical makeup, and some of us would like to say that its physical makeup makes it true that glass is fragile.[25] But notice that its physical makeup alone does not necessitate, *in the broadest metaphysical sense*, the truth that glass is fragile. For we also need, in addition, the actual laws of nature that make it such that when something has the physical makeup of glass, that thing is fragile. Now ought we to say that when we first

---

[25]The point here about fragility, construed qualitatively (rather than dispositionally), is merely illustrative. The same point can be made about men and mortality (in place of glass and fragility) if one prefers.

said that its physical makeup makes it true that glass is fragile, what we said was, strictly speaking, false and that instead we should hold that the laws of nature are, along with the physical makeup of glass, part of the truthmaker for ⟨Glass is fragile⟩? That seems absurd, for we should be able to at least remain neutral regarding what sort of things can count as truthmakers; in particular we should not have to posit higher-order entities like laws of nature as truthmakers. Truthmaker theory ought not commit us to there being extravagant truthmakers like *the laws of nature*.[26]

Mellor also gives up necessitarianism:

> If, as David [Armstrong] assumes, truthmakers must necessitate what they make true, it will take more than the truthmakers of 'Fa' and 'Fb' to make 'everything is F' true [assuming only a and b exist], since 'Fa & Fb' does not entail this, because it does not entail that there are no other particulars. But as 'there is no particular that is neither a nor b' is a negative truth, it needs no truthmaker. All it needs is that no truthmaker for its negation exists, i.e. that no particular other than a or b exists. So if a and b are indeed the only particulars, whatever makes 'Fa' and 'Fb' true will also make true 'everything is F', even though it will not necessitate it...In short, David's necessitation principle fails for generalizations, which are not entailed by the conjunction of all their instances, since that conjunction does not entail that there are no other instances. [Mel03, p. 214]

Mumford is yet another a rejecter:

> Armstrong's need for totality facts arises from his commitment to truthmaker necessitarianism, about which misgivings have already been expressed. If we do not need the facts to necessitate the truth, then it could be adequate for the truth of ⟨every x is F⟩ simply that ⟨every x is F⟩. We do not need an additional fact that these are all the x's, which then has to be understood as a kind of negative fact. [Mum05, p.268]

I conclude—and in good company, but that is no argument—that necessitarianism is false: truthmakers need not necessitate their truths. When we consider ordinary sentences and their truthmakers, eg. ⟨Fluffy is black⟩, necessitation holds trivially because the truthmaker is typically taken to be a trope and a trope necessitates a proposition because of how particularized an object it is. The blackness-of-Fluffy cannot exist unless Fluffy exists and is

---

[26]I do not deny there are laws of nature, in some sense of 'are'. I deny that laws of nature must be taken as objects to be quantified over. For Armstrong and others, laws of nature are relations between universals. But universals as such are objects that one might not wish to posit purely for the purposes of truthmaking, e.g. because universals are higher-order entities or because one does not believe in things such as relational tropes, and so on. Mellor makes the same point saying "Similarly with truths about David's beliefs, for example, that he is an Australian. For even physicalists will admit that it takes more than David's brain states to necessitate propositions about what he believes. It also takes laws linking his brain states to how he behaves, and perhaps his living in Australia and not in some 'twin Australia' elsewhere in the universe. Yet given all that, it is an innocuous abbreviation of physicalism to say that propositions about David's beliefs are made true by states of his brain" [Mel03, p. 214].

black. But not every proposition and its corresponding truthmakers are like this. In particular, universal propositions are not always like this. For one, they are never made true by tropes (though they may be made true by fusions of tropes). Second, a very simple and intuitive account of universals is the boolean account which, if true, implies the falsity of necessitarianism. Third, and most importantly, I see no good reason to generalize from some cases for which necessitation holds to all cases unless some *independent* grounds for necessitarianism are given. On the latter point, Armstrong attempts to provide such independent grounds with the following argument:

> Suppose that a suggested truthmaker T for a certain truth p fails to necessitate that truth. There will then be at least the possibility that T should exist and yet the proposition p not be true. This strongly suggests that there ought to be some further condition that must be satisfied in order for p to be true. This condition must either be the existence of a further entity, U, or a further truth, q. In the first of these cases, T + U would appear to be the true and necessitating truthmaker for p. (If U does not necessitate, then the same question raised about T can be raised again about U.) In the second case, q either has a truthmaker, V, or it does not. Given that q has a truthmaker, then the T + U case is reproduced. Suppose q lacks a truthmaker, then there are truths without truthmakers. The truth q will hang ontologically in the same sort of way that Ryle left dispositional truths hanging. [Arm04, pp. 6–7]

However, two main worries immediately arise. The first is that the argument shows at most that (assuming Maximalism) for every truth there is a truthmaker that necessitates that truth. It does not show that every truthmaker necessitates every truth it is a truthmaker for! The second is that the argument begs the question. For why should we think that because T does not necessitate the truth of p that "there ought to be some further condition that must be satisfied in order for p to be true". That is precisely what the anti-necessitarian rejects.

If one has any affinity to familiar model theory for quantificational languages—e.g. as getting *roughly* correct the truth conditions for a (fragment of a) regimented natural language—she should also not find necessetarianism attractive. For such a theory interprets (relative to a model) a universal sentence as the conjunction (up to logical equivalence) of its instances. Moreover, model theory may be viewed as modeling a correspondence theory of truth (more precisely, of satisfaction), where the truth of sentences depends on elements of a given domain of individuals. Indeed one may view sequences of the domain as truthmakers for sentences. The intuitive appeal of identifying universal propositions with conjunctions provides another mark against the necessitarian.

(A legitimate problem is *defining*, in a given language L, a universal *statement*, *not proposition*, $\forall x A$ as the conjunction (up to logical equivalence) of its instances. For in this case there are models M that satisfy each instance $A(t)$ in L (for t a closed term of L) without satisfying $\forall x A$. That is why an omega rule in a usual first-order language is not sound over the class of all models. But that is not a problem for identifying universal *propositions* with the conjunction of their instances.)

Intensional propositions

If every intensional operation can be given an analysis in terms of restricted quantification over worlds (or information states, situations, etc.) then intensional propositions are effectively quantified propositions, and we already have an account that tells us when such propositions are positive. We would thereby get an account of positivity for intensional propositions for free, so to speak. Numerous have and do hold that intensionality is reducible to the extensional via restricted quantification. One might, however, deny the prospects of extensional reduction–though I have not seen any compelling arguments establishing this (but for an interesting argument to the contrary, see [RM77])—and argue that the best we might hope for is a laundry list of positive property-forming operations. This seems to me unsatisfactory for two reasons. The first is that it denies a uniform treatment of intensionality in terms of quantification without giving a general account of intensionality in its place—assuming no such account has been so given. The second is that it makes difficult a simplified, and to me highly satisfactory, theoretical analysis of the class of positive propositions, a theory which has fruitful applications inside and outside of truthmaker theory.

Suppose then that intensionality can be given an extensional analysis. Then

Intensional positivity. An intensional proposition $\langle A \rangle$ is positive iff $\langle \phi_A \rangle$ is, where ϕ_A is the extensional reduct of A.

As a concrete case, suppose the intended semantics of the **S4**-$\square$ is given by its usual Kripke semantics. Then whether a modal proposition $\langle \square A \rangle$ is positive depends on whether $\langle \forall w(R(@, w) \to A(w) \rangle$ is positive, where @ is the actual world, $A(x)$ means that x is true at A (e.g. in the "standard model" of **S5** consisting of all possible worlds) and R is a preorder on worlds (again, e.g., from the standard model).

There is much more to say here that will have to be left for another occasion. However I should say, at least, that what I take to be the standard model is ersatzist—possible worlds are mere constructions whose transitive closures are such that all of their urelements actually exist. Hence I reject any serious form of possibilism. Notice that this must be so, since even modal statements are made true by the maximal truthmaker which is itself the mereological sum of all *actual* things: modal truths are ultimately grounded in the non-modal.

Positivity in general

We now have all the essential ingredients for a complete analysis of positivity. To summarize, we have the following:

- Every atomic proposition is positive;

- $\langle \neg A \rangle$ is positive iff $\langle A \rangle$ is negative (i.e. not positive);

- $\langle A \wedge B \rangle$ is positive iff both $\langle A \rangle$, $\langle B \rangle$ are positive;

- $\langle \forall x A \rangle$ is positive iff each of its instances $\langle A(t) \rangle$ is positive;[27]

---

[27]There are no restrictions on the language here so that e.g. objects may name themselves. We thus avoid any loss of generality by not having enough names in the language.

- $\langle \nabla A \rangle$ is positive iff its extensional reduct $\langle \phi_A \rangle$ is positive, where ∇ is an intensional operator.

Given the usual assumptions about (classical) duality and definability (e.g. of quantifiers and modal operators), the definition of positivity above is extremely general in character.

The following restricted version of Maximalism is endorsed on the present account:

Every positive truth has a truthmaker.

On the other hand, a proposition lacks a truthmaker just in case it is negative. But not all propositions that lack truthmakers are false, so being false is not to be equated with lacking a truthmaker.

5.8   Final remarks

I have motivated a truthmaker theory whereby only positive propositions have truthmakers and being made true in the grounded sense is equivalent to being made true by the maximal truthmaker, and being made true in the ungrounded sense is equivalent to lacking a truthmaker. Thus if truth values are objects, truth as the maximal truthmaker is the only truth value. The redundancy of falsity was always clear from a formal point of view, being there merely to ensure that assignment functions be total rather than partial, so the present philosophical rejection of falsity further serves as a reduction of primitives both conceptually and ontologically.

As a methodological platitude of theorizing, one ought not invoke for those purposes entities whose natures are mysterious or suspect. In the course of history, the invocation of truth values without question was a gross violation of this platitude. Truth being something as wholly unmysterious as the maximal truthmakers allows one to invoke truth-as-object without any longer violating this platitude.

Chapter 6

The definability of negation as impossibility

Abstract

A strong notion of negation is that of impossibility. Intuitionistic negation when modally construed is an example of negation as impossibility. Negation as "absolute" impossibility, where possibility is taken in its broadest sense, is not easy to define within the confines of the usual modal languages. We investigate some languages that are capable of defining negation as impossibility when necessity corresponds to the global modality. The languages considered include counterfactual operators with propositional quantification and a bimodal language including a modality and its complementary. Soundness along with some preservation and translation results for the bimodal language are given.

Negation as a modal operator has been studied extensively in [Dos86], [Dos99] and [Dun05]. The present perspective on the situation is much different. Its emphasis is on the definability of negation as "absolute impossibility", that is, impossibility in its strongest sense as falsity in every point of a model, not just those that are accessible from a given point of evaluation (unless, of course, accessibility is universal).

Typically an interest in negation as impossibility, regarded from a semantic perspective, is coupled with an equally strong interest in obtaining a logic characterized by the semantics that has certain niceties such as validating formulae regarded to be "characteristic" of negation or some other connectives.[1] One such formula might be $(A \to B) \to (\neg B \to \neg A)$, its rule form referred to in chapter 2 as IContra (Intuitionistic Contraposition). The interest here lies not as much in the niceties as the mere definability of impossibility in various languages that arise commonly in the philosophical literature. Two such languages to be investigated include a counterfactual operator and propositional quantification, and another with a modal operator and its complementary. The

---

[1]A number of such formulae are listed and discussed in chapter 2.

complementary modality has been investigated in the literature under various names; in [GPT87] it is referred to as a "sufficiency operator", in [Hum83] and [Gor90] as a "complementary" modality, and in [Gol74] and [Dun93] as a negation. Since these languages will also include boolean negation, the problem of defining negation as impossibility is reduced to defining the global modality whose truth conditions are given by: 'It is necessary that A' if and only if 'It is true in every point of the model that A'.

The chapter proceeds as follows. We first start with some preliminaries followed by a brief discussion of the logic of negation as impossibility. Then two counterfactual semantics with propositional quantification are investigated and the global modality is defined by placing certain restrictions on the system of spheres and the selection function. Finally we look at a language with a modality and its complementary modality. Soundness and some preservation and translation results are given for this language.

6.1 Preliminaries

We start by getting some preliminaries out of the way. Let L be the basic propositional language consisting of a set $Prop$ of propositional letters $p_1, \ldots$, the (atomic) falsum $\perp$, implication $\to$ and the necessity operator $\Box$ with other connectives defined as usual. $L(\otimes_1, \ldots, \otimes_n)$ is L augmented with the n operators $\otimes_1, \ldots, \otimes_n$, i.e. $L(\otimes_1 \cdots \otimes_n) = L \cup \{\otimes_1, \ldots, \otimes_n\}$. We may write $L(\Delta)$ for short when $\Delta = \{\otimes_1, \ldots, \otimes_n\}$. We call an n-ary modal operator $\otimes$ (semantically) *Kripkean* if $M, a \models \otimes(A_1, \ldots, A_n)$ iff $\forall b_1, \ldots, b_n$ s.t. $R_{\otimes}ab_1 \ldots b_n$, $M, b_i \models A_i$ $(0 \leq i \leq n)$.[2] Uppercase Latin letters A, B, etc. are metavariables ranging over sentences of a given language. Expressions are used autonomously unless quote marks would aid in readability.

A *frame* F is a pair (W, R) consisting of a non-empty set W of worlds and $R \subseteq W \times W$ a binary relation W. A *model* $M = (F, V)$ (or equivalently (W, R, V)) *based on* the frame F consists of a frame together with an assignment $V : Prop \to \wp(W)$ of subsets of W (propositions) to propositional letters. The *truth set* $\|A\|^M = \{a \in W : M, a \models A\}$ of a formula A relative to a model M is the set of worlds at which the formula is true. Customarily we drop the superscript indicating the model when it is clear. It will also be convenient to talk about "A-worlds" (relative to a model) as worlds that are members of the truth set of A.

One of the central concepts that concerns this chapter is the *definability* of frame properties. For example, if a frame has an alternative relation R that is universal (i.e. $R = W \times W$), we call the frame 'universal'. This is especially helpful when there is but one alternative relation. In the general case, when frames have any number of alternative relations, it might be less helpful to speak this way. (If a frame has a transitive alternative relation and an non-transitive one, calling the frame 'transitive' and also 'non-transitive' is awkward, though perhaps not as awkward as calling it 'transitive and non-transitive' in the same breath. Below we will be working with frames with two alternative relations.) The following gives a precise formulation of definability.

---

[2] These sorts of operators are sometimes called 'universal' in the literature. This nomenclature would no doubt be confusing for our purposes.

Definition 6.1.1. We say a set Γ of $L(\otimes_1 \cdots \otimes_n)$-formulas defines a property P on a relation R_i $(1 \leq i \leq n)$ with respect to the class $\mathcal{K}$ of frames $F = (W, R_1, \ldots, R_n)$ iff for any frame $F \in \mathcal{K}$

$$F \models \Gamma \text{ iff } R_i \text{ has } P.$$

Now consider the standard modal language L which fixes for $\square$ the satisfaction condition

(6.1) $M, a \models \square A$ iff $\forall b (Rab \Rightarrow M, b \models A),$

for a, b worlds from M. Then the following holds.

Proposition 6.1.2. *Universality is not definable in L, i.e. there is no set Γ of L-formulas such that for all frames $F = (W, R)$, $R = W \times W$ iff $F \models \Gamma$.*

Proof. Suppose there were a set Γ in L that defines universality. Take any two disjoint frames $F_1 = (W_1, R_1)$ and $F_2 = (W_2, R_2)$ with each R_i universal. We have $F_i \models \Gamma$ $(i = 1, 2)$ by supposition. By the preservation of formulas under taking disjoint unions of frames (see [BdRV02, p. 53] for details), $\biguplus_{i \leq 2} F_i \models \Gamma$ (where $\biguplus_{i \leq 2} F_i$ is the disjoint union of the F_i). But the alternative relation of $\biguplus_{i \leq 2} F_i$ is not universal which contradicts our hypothesis. $\square$

This tells us an important fact about L: it is unable to discriminate structures for necessity from structures for other modalities such as epistemic and nomological ones which themselves are often interpreted by equivalence frames, as e.g. in the famous **S5** interpretation of 'S knows that'.[3] In the following section we introduce a quantified propositional language with a counterfactual conditional in the spirit of Lewis and Stalnaker and determine which properties of the frames for the language are required for defining the global modality. In certain cases they turn out reasonably minimal.

In what follows we denote the universal global modality by $\blacksquare$ whose truth conditions are:

(6.2) $M, a \models \blacksquare A$ iff $\forall b (M, b \models A).$

Its existential dual $\blacklozenge$ is defined by

(6.3) $M, a \models \blacklozenge A$ iff $\exists b (M, b \models A).$

6.2   Negation as impossibility

The McKinsey-Tarski translation of the language of intuitionistic propositional logic (**IPC**) into the language of **S4** is a natural one.[4] It is also one according to which intuitionistic negation $\neg$ is translated as the impossibility modality definable by $\square \sim$ where $\sim$ is boolean negation.[5] Kripke semantics for **IPC** makes use of the same structures as those for **S4**, viz. the class of preorders, but differs from standard modal logics in imposing the following Heredity constraint:

---

[3]In [HBW08] it is claimed that "the logical system **S5** is by far the most popular and accepted epistemic logic". Other modalities plausibly interpreted by equivalence relations are "It is a law that", "It is metaphysically necessary that", "It is correctly assertible that", and so on.

[4]The reader is referred to [MT48, p. 13] for details.

[5]To be precise, the truth conditions for boolean negation are: $M, a \models \sim A$ iff $M, a \not\models A$.

(Heredity) If $M, a \models p$ and $a \leq b$ then $M, b \models p$.

The condition, extendable to arbitrary formulae, is equivalent to requiring that the valuations assign only upsets to atoms.

Suppose we pair the class of structures down to the universal frames and drop Heredity. Taking the class of arguments valid over such a semantics then gives us a logic that is a close relative to intuitionistic logic except that negation now has the interpretation of absolute impossibility. (Notice that if we did not drop Heredity the resulting logic would be classical, for truth at a point would be equivalent to truth at every point, i.e. models would contain pairwise modally equivalent points.) And the logics really are close, especially with respect to negation. For instance, all the laws characteristic of intuitionistic negation, in the sense of either holding (e.g. $A \to \neg\neg A$) or not holding (e.g. $\neg\neg A \to A$), either hold or not for negation as **S5** impossibility (assuming implication is strict) depending on what is the case for **IPC**. Call this logic *strict* **S5** or **SS5** for short. Then what we have just noticed is that a characteristic negation law A holds in **IPC** iff it holds in **SS5**. The notable difference between the two logics' characteristic principles do not concern negation. For example **ICP** endorses

- $A \to (B \to C)$;

- $A \to (B \to (A \wedge B))$;

- $(A \to (B \to C)) \to (B \to (A \to C))$

while **SS5** does not.

One might expect the negation axiom $A \to \neg\neg A$ (DNI) to separate the two logics since it translates (on the McKinsey-Tarski translation) into the *Brouwerische* axiom $B := A \to \Box\Diamond A$ characteristic of symmetry, a condition which is valid only over the frames for **SS5** and not **IPC**. But it is of course intuitionistically valid, holding in virtue of Heredity (and reflexivity) rather than symmetry. We are left hard pressed to find any difference concerning negation between the two that is not merely an instance of a difference concerning essentially implication and connectives other than negation.

So from an *internal* perspective (i.e. from within the logic, looking at the logical truths) we obtain a kind of constructive negation by just taking the frames for **S5** and giving "intuitionistic" truth conditions to the connectives (while not imposing Heredity).[6] (From an *external* perspective it might be thought that certain properties of the logic such as the disjunction property hold in order that the logic be deemed 'constructive' in any suitable sense. We put this issue to the side.) In particular the negation of **SS5**, like intuitionistic negation, does not validate the following:

- $\neg(A \wedge B) \to \neg(A \vee \neg B)$;

- $\neg\neg A \to A$;

- $A \vee \neg A$;

---

[6]We might instead think of the intuitionistic connectives as taking on the Lewis truth conditions for strict implication and negation defined as implication to absurdity, though Lewis never thought of negation this way. For him, negation was thoroughly boolean.

- $(\neg A \to \bot) \to A$;

- $(\neg A \to \neg B) \to (B \to A)$;

- $(A \to B) \lor (A \to \neg B)$

each of which has a thoroughly non-constructivist interpretation.

6.3 Extended modal languages

We first look at definability of the global modality in conditional languages, starting with frames based on (i) selection functions, and (ii) systems of spheres. Frames of the first type are due to Stalnaker [Sta68] and those of the second to Lewis [Lew73]. Depending on the semantic constraints one accepts, definability according to these two different kinds of structures is not obtained in parallel fashion. We give two versions of the selection function semantics. The first is a generalization of Stalnaker's original account, and the second is essentially the original. In a sense the original seems more natural, but only under the assumption that selection functions are total (for every world there is at least one "closest" one) and have as their codomain the set of worlds rather than its powerset (for every world there is a unique closest one).

The next language we look at in section 6.3 involves a sufficiency (or "window") modality. Such modalized statements express that a world w' is an alternative of w provided some formula holds at w—i.e. they provide sufficient conditions for alternativeness (or "accessibility"). Besides its use in perp or (in-) compatibility semantics as a negation operation, the modality has other philosophical applications as well. For instance the holding of a sentence in w (say, 'There is a counterpart in w' of an individual in w') might imply that some other world w' is an alternative of w.[7]

The modality occurs naturally in the setting of boolean modal operators where it was originally introduced. Details may be found in [BdRV02, ch. 7].

Conditional logic and propositional quantification

One natural rendering of 'It is necessary that A' is 'For any proposition B, had B been true, A would have been true too', or briefly, 'A would hold no matter what holds'.[8] Such a rendering of 'It is necessary that A' involves propositional quantification, initially introduced by Fine [Fin70]. Logics with propositional quantification are often referred to as second-order propositional modal logic (**SOPML**) and it was shown in [KT97] that **SOPML** based on **S4.2** or weaker modalities is equal in model-theoretic strength to full second-order predicate logic (**SOL**). Fine showed these logics to be incomplete, and in [Gar01] Garson shows a similar incompleteness result for second-order modal arithmetic based on **S4.3** or weaker modalities.

---

[7]Less philosophically, consider nominals again. If i is a nominal naming world w_i, then the truth of $\Diamond i$ at w implies that w_i is an alternative of w. Clearly this is a general principle which holds for any i.

[8]What a proposition is will be left intentionally vague, but one way to think of them is as sets of possible worlds (or as some might say, in the "UCLA" way). Then each member of a proposition is a world in which that proposition holds. Sometimes worlds themselves are called 'propositions', being maximally consistent descriptions of the way the actual world might be.

The impressive strength of **SOPML** is not surprising as quantifiers have their expected interpretations: $\forall p_i A$ is true at a world just in case for every subset X of worlds, $A[V(p_i) \mapsto X]$ (i.e. the result of interpreting p_i as X) is true. Propositional quantifiers in a modal setting thus allow us to quantify over arbitrary subsets of a possibly infinite set of worlds. The similarities between **SOPML** and **SOL** are further brought out in translation theory. For example, [Fin70] translates second-order arithmetic into propositionally quantified **S4.2** thereby showing the former to be incomplete.

If **SOPML** is so powerful, why can we not define the global modality in the basic modal language L extended to include propositional quantifiers? The reason is quite simple: propositional quantification does not give us access to inaccessible worlds. Consider one plausible definition of $\blacksquare A$: $\forall p_i \Box (p_i \to A)$.[9] The propositional quantifier does nothing in allowing us to extend our reach to inaccessible worlds, though it allows us to say other interesting things not expressible in L. We will see that things are different when we tinker with this definition by removing the box and replacing the material arrow with a conditional one.

While we are defining the global modality in terms of counterfactuals, it is interesting to note that things may be done the other way around as done in [vBRvO06]. So in a sense the global modality and certain counterfactuals are interdefinable. However, the counterfactuals definable in the presence of the global modality, in the sense of [vBRvO06], are not the sort that figure in the mainstream philosophical literature, and perhaps for good reason. The counterfactual of [vBRvO06] interprets 'Were A, B' as 'For any world w in which A holds, there is a world w' better than w at which A strictly implies B', which seems to us to be a poor interpretation of a "would"-counterfactual for the following reason. The fact that for every A-world there is a better B-world at which A strictly implies B does not rule out there also being a better B-world at which the strict implication from A to $\neg B$ holds. But then while it is true to say that the consequent would have obtained had the antecedent, the same can be said of the negation of the consequent, and standardly 'Would A, B' and 'Would A, $\neg B$' are thought to be mutually inconsistent.

One may define the counterfactuals of more interesting logics by the methods of [Lew73], where e.g. a Lewis counterfactual is defined in terms of propositional quantification and a "sphericality" operator. One may then go on to define the global modality in the way set out in the following section, in a less direct fashion. Obviously there are a number of ways to define the global modality and I have merely chosen to focus on three I find interesting.

Selection functions

We work in the language $L(\forall, >)$[10] where $>$ is a binary conditional operator. A frame F for this language is a pair (W, f) where $f : W \times \wp(W) \to \wp(W)$ is a *selection function* (*s-function* for short) mapping propositions to world-proposition pairs. We call such a frame an *s-frame*. An (s-) model is a pair

---

[9]In the relevant systems we have $\forall p_i \Box (p_i \to A) \to \Box \forall p_i (p_i \to A)$, so it suffices to make our point with the aforementioned formula rather than $\Box \forall p_i (p_i \to A)$.

[10]Technically the base propositional language L specified in section 6.1 has no variables for binding. We assume here the obvious modifications to L, viz. that it contain denumerably many variables as well as constants.

(F, V) where V is as usual. Roughly, one thinks of the antecedent $f(a, \|A\|)$ of a counterfactual as the collection of A-worlds closest to a, which intuitively expresses the proposition that A is entertainable with respect to a. A counterfactual is to be true exactly when this proposition *implies* the consequent, in the sense that A implies B in a model when $M \models A \to B$ holds, or equivalently $\|A\| \subseteq \|B\|$. In effect, the s-function picks out the set of worlds that are most similar to the world of evaluation with respect to the antecedent.

We employ a nearly minimal conditional semantics which places the following sole restriction on f:

(6.4) $\qquad$ if X is a singleton then for all $b, f(b, X) = X$

for $a, b \in W$ and $X \subseteq W$. The condition by itself is relatively weak. It says that if A is true at exactly one world, then that world must be the closest A-world to any given world. The condition is plausible enough and is reminiscent of Stalnaker's uniqueness assumption that a proposition is possible (in the global sense) only if there is a unique most similar world making it true.[11] But notice that we have not placed the uniqueness condition on the s-function generally. This is what makes the present condition comparatively weak, but also less motivated. The purpose of first presenting the semantics this way is to see just what is needed in the more general setting. Later we briefly discuss dropping this restriction in favor of others that get frequently mentioned in the literature, some combinations of which make quantification unneeded (and are hence jointly *stronger* than our present restriction).

The satisfaction clause for propositional quantification is as follows. Let $M = (W, f, V)$ be an (s-) model. Then

$$M, a \models \forall p_i A \text{ iff } (W, f, V'), a \models A$$

for every V' like V except at most in what it assigns to p_i. We call V' a p_i-variant of V. For the conditional arrow we define

$$M, a \models A > B \text{ iff } f(a, \|A\|) \subseteq \|B\|.$$

We read $A > B$ as "Were A, B" in accordance with our intended counterfactual interpretation of the conditional.[12]

Since we have quantification over propositions, in particular we have quantification over singleton subsets $\theta \subseteq W$. But if every such θ implies A, by the restriction on f, that means every world from the model yields A—i.e. A is *globally* true in the model. Thus quantification coupled with conditional implication gives us enough for the following definition of the global modality:

$$M = (W, f, V), a \models \forall p_i(p_i > A) \text{ iff } M, a \models \blacksquare A.$$

We may reason through the definition as follows. The left side holds iff $M' = (W, f, V'), a \models p_i > A$ for every p_i-variant V' of V, iff $f(a, \|p_i\|^{M'}) \subseteq \|A\|^{M'}$ iff $f(a, \{b\}) \subseteq \|A\|^M$ for all $b \in W$, iff $\|A\|^M = W$ iff $M, a \models \blacksquare A$.

---

[11] There may be cases where we want to say that all worlds are equally remote with respect to some other world. Lewis allows this on both of his semantics, one involving systems of spheres and the other comparative possibility. Our restriction prohibits these sorts of cases.

[12] The reading is intended merely as an informal gloss. If f lacks certain properties then the formulas validated over frames based on f will not contain certain other formulas thought characteristic of a counterfactual conditional. That is true as it stands here, since f is restricted only by (6.4).

It is clear this reasoning relies essentially on condition (6.4), and there does not appear to be any *weaker* restriction we can place in its stead. Here are three obvious suggestions:

1. $f(a, X) \subseteq X$;

2. if $a \in X$ then $a \in f(a, X)$;

3. if $X \neq \emptyset$ then $f(a, X) \neq \emptyset$.

The first of these is called '(id)' by Chellas in [Che75, p. 142] and corresponds to (is defined by) the identity schema $A > A$. The second also appears in [Che75] as '(mp)' and corresponds to $(A > B) \to (A \to B)$, an expression of modus ponens for $>$. Finally, the third condition requires that, unless a proposition is impossible, there must be some similar world at which it holds. This is similar to our initial singleton restriction (6.4), and indeed it implies (6.4) under the plausible assumption that the first condition holds. However, the first and last conditions make quantification redundant; we can instead adopt Lewis' original formulation (in [Lew73]) of global necessity as $\neg A > \bot$.

It is also interesting to note that, given the global modality, we can define the third condition by $\blacklozenge A \to \neg(A > \bot)$. A simple semantical argument suffices. Suppose the third condition holds and that $M, a \models \blacklozenge A$. Then $\|A\| \neq \emptyset$, so by the condition, $f(a, \|A\|) \neq \emptyset$. Whence $f(a, \|A\|) \not\subseteq \|\bot\|$, so $M, a \not\models A > \bot$ and thus $M, a \models \neg(A > \bot)$.

For the other direction, suppose $F \models \blacklozenge A \to \neg(A > \bot)$ and $X \neq \emptyset$. Let $V(p) = X$. Then $(F, V), a \models \blacklozenge p$, so $(F, V), a \models \neg(p > \bot)$. Whence $f(a, \|p\|) \not\subseteq \|\bot\|$ iff $f(a, \|p\|) \neq \emptyset$ as desired.

We have taken f to be an s-function from worlds to sets of worlds. But suppose we take f to be a *total* function from world-proposition pairs to worlds satisfying the condition $f(a, X) \in X$. That is, the closest world with respect to the counterfactual is an antecedent-world. This is essentially Stalnaker's original semantics. For X a singleton $\{b\}$ the condition reduces to $f(a, \{b\}) = b$. The semantics (and the notions of s-frame and s-model) for the counterfactual is adapted straightforwardly to

$$M, a \models A > B \text{ iff } f(a, \|A\|) \models B.$$

Again $\forall p_i (p_i > A)$ defines the global modality by the totality and uniqueness of f.

Systems of spheres

We briefly discuss the definability of the global modality within the conditional semantics of Lewis, ingeniously developed in [Lew73]. Lewis was aware of a global modality (which he thought of as the broadest logical modality) and that the satisfaction of a particular semantic constraint ensures that the outer modality is global. But he says little more about it. To say that a sentence is necessary within a particular system of spheres, however, seems to have little in common with saying that it is necessary *simpliciter* (i.e. in the global sense we intend). Lewis seems to agree; he states "a sentence is *necessary*, *possible*, or *impossible* iff it is true at all worlds, at some, or at none" [Lew76, pp. 299]. For a proposition is hardly necessary in the broadest sense if it is true at some

worlds and false at others, and it is precisely this possibility that is permitted by the semantical system of [Lew73]. Indeed, the global modality is not definable in this system; i.e. no formula in the purely counterfactual language defines the aforementioned semantic constraint. We can get it, as we did above for s-frames, by adding propositional quantification. We go straight to the details.

Our language is $L(\forall, >)$, a frame for which is a pair (W, f) where f maps a (centered-) system of spheres $\mathcal{S}_a \subseteq \wp(W)$ to each world a. (A model for the language is a pair (F, V) where F is a frame for the language and V is as usual.) We call such frames *sphere* frames. A system of spheres $\mathcal{S}_a$ is a family of subsets of W closed under the following conditions:

1. $\{a\} \in \mathcal{S}_a$ ($\mathcal{S}_a$ is *centered* on a);

2. for each $X, Y \in \mathcal{S}_a$, either $X \subseteq Y$ or $Y \subseteq X$ ($\mathcal{S}_a$ is $\subseteq$-*dichotomous*)[13];

3. $\mathcal{S}_a$ is closed under arbitrary (i) unions and (ii) non-empty intersections.

Each member of $\mathcal{S}_a$ is called a 'sphere', each pair of which is $\subseteq$-comparable as stated in the second condition. If $\emptyset \notin \mathcal{S}_a$ then the *innermost* sphere is its centering, $\{a\}$; otherwise it is $\emptyset$. The reason for this is that the innermost sphere is defined as $\bigcap \mathcal{S}_a$, and since systems are centered, either $\bigcap \mathcal{S}_a$ is $\{a\}$ or it is $\emptyset$. Intuitively, the innermost sphere contains worlds that are the most similar to a, and surely a is more similar to itself than any other world.[14]

The satisfaction condition for counterfactuals is disjunctive, one case covers the vacuous case and the other what Lewis calls the "principal" case:

$$M, a \models A > B \text{ iff } \left\{ \begin{array}{ll} \text{either} & \bigcup \mathcal{S}_a \cap \|A\| = \emptyset \\ \text{or} & \exists S \in \mathcal{S}_a : S \cap \|A\| \neq \emptyset \wedge S \subseteq \|A \to B\|. \end{array} \right.$$

A counterfactual $A > B$ is vacuously true if its antecedent is not true anywhere in the system of spheres—in Lewis' words, it is not "entertainable"—otherwise it is made true by some A-permitting sphere (i.e. a sphere containing some A-world) in which the material implication $A \to B$ holds throughout. By dichotomy such a sphere will be "minimal".

Lewis provides the following two equivalent definitions of a necessity modality:

$$\Box A := \begin{array}{l} \neg A > \bot \\ \neg A > A. \end{array}$$

These are analogous to necessity defined in terms of strict implication. Informally a proposition is necessary if it is true at every world in the system and it is possible ($\Diamond A \leftrightarrow \neg \Box \neg A$) if there is a world in the system at which it is true. Since conditional formulas are evaluated with respect to both worlds and systems of spheres, the above modality is inherently restricted to the relevant

---

[13]Lewis calls this condition *nestedness*.

[14]If we took worlds to be e.g. maximal consistent sets, then there are no indistinguishable (i.e. modally equivalent) worlds: i.e. there are no two worlds making precisely the same formulas true. Then clearly each world is most similar to itself. On the present account, however, it is possible that two worlds be indistinguishable, in which case, each is equally similar to the other and to themselves. Even in this case Lewis apparently thought they are not completely indistinguishable: if we call one 's' and the other 't', then only the former has the property of being identical to s. Despite this, Lewis remained neutral on whether there really are distinct yet indistinguishable worlds.

systems of evaluation and so may not extend in its reach to every world unless each system of spheres does. Precisely, $>$ extends to every world iff $\bigcup \mathcal{S}_a = W$ for each $a \in W$. Lewis calls frames satisfying this constraint *universal*. It is easy to verify that the above definitions for $\square$ are equivalent to those for $\blacksquare$ over universal frames.

The axioms provided in [Lew73, p. 121] that Lewis cites as "characteristic axioms" for universality are characteristic in a weak way: they are valid on any universal frame, though the converse, that their frame validity ensures universality, need not hold. For instance, there are frames validating schemas T, 4 and 5 (defined in terms of $>$) that are not universal. Again, this is due to the fact that the evaluation of conditionals does not extend to what we might call the most 'remote' worlds, worlds that lie beyond any sphere in the system.

The condition

$$\forall a, b \in W, \bigcup \mathcal{S}_a = \bigcup \mathcal{S}_b,$$

Lewis calls *uniformity*. Together with the assumption that frames are centered, as we have presented them here, uniformity entails universality. For if frames are centered (and hence *totally reflexive* in the jargon of Lewis, i.e. $\forall a \in W, a \in \bigcup \mathcal{S}_a$), then every world is contained in some sphere, viz. the sphere with that world as its center. And if each system of spheres for each world shares the same stock of worlds, then that stock must include all worlds. If we add quantification to the language we come close to defining universality by

$$(6.5) \qquad\qquad\qquad \forall p_i \Diamond p_i,$$

which states that every proposition is "possible", including all singletons, and hence that for every $X \subseteq W$ there is a b in each system such that $b \in X$. Clearly this is false since there is a valuation assigning p_i the empty set, and thus under this valuation p_i is true at no world in any system of spheres. But then $\forall p_i \Diamond p_i$ is a full-fledged contradiction, equivalent to $\forall p_i p_i$ and to $\bot$.

The above problem can be seen as one among many of a general sort for conditional logics that concerns how to handle impossible antecedents. In Lewis' system they are vacuously true. In his modified semantics which requires that there be an antecedent-permitting sphere in order that a conditional be true, they are vacuously false. But in either case, there is good reason to think that not all conditionals with impossible antecedents are semantically equal. For example 'If there were a largest prime, not every integer would have a prime successor' rings true, while 'If there were a largest prime, there would not be a largest prime' sounds necessarily false. To accommodate impossible antecedents in a more faithful fashion we adopt "impossible" worlds. They have been used in e.g. [ST70] as a mere technical device for this end.

Impossible worlds are intended to be worlds at which anything goes. In particular they make $\bot$ "true".[15] They may be devised so that not all formulas are true, just some impossible ones are. This allows a more fine-grained distinction between the meaning of counterfactuals with impossible antecedents. For simplicity we let impossible worlds make everything true. If we outfit

---

[15]If this seems philosophically distasteful, another strategy is to define $\bot$ by a sentence of the form $A \wedge {\sim} A$ where $\sim$ is a negation operation such that A and ${\sim} A$ may be simultaneously true (forced) at a world as in the semantics of certain substructural logics such as the relevance logic **R**. Accordingly we can do away with impossible worlds in which the truth conditions of certain formulas may be non-compositional. Such is the case with λ mentioned below.

each system of spheres $\mathcal{S}_a$ with precisely one impossible world λ, then (6.5) defines uniformity. The closest worlds to any given world with respect to an impossible antecedent will be λ, and so when $V(p_i) = \emptyset$ we have its possibility following from its truth at λ, i.e. $(F, V), a \models \Diamond p_i$ follows from $(F, V), \lambda \models p_i$ and $\lambda \in \bigcup \mathcal{S}_a$. The global modality is then defined by

$$\blacksquare A := \forall p_i \Diamond p_i \wedge \Box A.$$

Axiomatics

The *basic conditional logic*, denoted by **CK** in accordance with [Che75], is defined as the class of all formulas valid on every s-frame $F = (W, f)$, where f is not required to satisfy any restriction.[16] It is axiomatized by

1. all classical propositional tautologies;

2. $(A > (B \wedge \theta)) \to ((A \to B) \wedge (A \to \theta))$;

3. $((A > B) \wedge (A \to \theta)) \to (A > (B \wedge \theta))$;

4. $A \to \top$,

and the rules

1. $A, A \to B \vdash B$;

2. $A \leftrightarrow B \vdash (A > \theta) \leftrightarrow (B > \theta)$;

3. $(A_1 \wedge \cdots \wedge A_n) \to A \vdash ((B > A_1) \wedge \cdots \wedge (B > A_n)) \to (B > A)$.

The *second-order basic conditional logic*, **SOCK**, is **CK** together with the quantificational axioms

1. $\forall p_i(A \to B(p_i)) \to (A \to \forall p_i(B(p_i)))$ (p_i not free in A);

2. $\forall p_i A(p_i) \to A(B)$,

and the generalization rule

1. from $\vdash A(p_i)$ infer $\vdash \forall p_i A(p_i)$.

It is an open problem whether there is a recursive axiomatization of the class of $L(\forall, >)$-formulas valid on all s-frames satisfying our singleton restriction (6.4). One problem is due to the inability to express over arbitrary structures whether some formula is true in precisely one world. Augmenting the language further with a "difference" operator D (see e.g. [BdRV02] and [dR92]) defined by

$$M, a \models DA \text{ iff } \exists b \neq a, M, b \models A$$

would give us just that ability, but it also smuggles in with it a definition of the global modality as $A \wedge \neg D \neg A$. Ideally we would like to define singleton propositions by quantification and the conditional arrow alone and to leave the smuggling as a means for obtaining other desirable goods.

---

[16]We make no distinction between a logic as (i) a class of formulas **L** and (ii) a deduction system which generates precisely **L**.

It is worth mentioning an axiomatization of the logic of s-frames satisfying the first and third alternative conditions mentioned previously on page 107. Recall that together they give us the global modality $\blacksquare A := \neg A > \bot$ without propositional quantification.

An axiomatization is given by **CK** plus

1. $A > A$;

2. $(A > \bot) \to \blacksquare \neg A$.

The resulting axiomatization is sound and complete with respect to the class of s-frames satisfying the given conditions. In fact, it is easy to verify these axioms define their respective corresponding frame properties.

The sufficiency modality

We now look at the sufficiency modality. A completeness proof of an axiomatization of this modality can be found in [Hum87]. For definability results, [Gor90] provides a rich source.

The following modality, which we denote by $\boxplus$, is referred to in [BdRV02] as a "window" operator and in other places as a "sufficiency" and as a "complementary" operator. Goldblatt [Gol74] popularized the operator as a negation $\sim$ defined in [Dun93] by

$$M, a \models \sim A \text{ iff } \forall b(M, b \models A \Rightarrow b \bot a),$$

for $\bot$ the relation of incompatibility. If incompatibility is asymmetric then there arises both a left- and right-negation.

The satisfaction clause for $\boxplus$ is unsurprisingly

(6.6) $$M, a \models \boxplus A \text{ iff } \forall b(M, b \models A \Rightarrow Rab).$$

Loosely, a formula $\boxplus A$ is true at a world a if and only if a can see every A-world. It may also be viewed as providing sufficient conditions for R-accessibility, whereas the standard Kripkean clauses provide necessary conditions. In [Hum87] a primitive modality which is equivalent to $\boxplus \neg A \wedge \Box A$ is studied as a providing a logic of "all and only". With this modality in hand, it is easy to define universality and the reader should verify that $\boxplus \top$ does the trick.

The truth of $\boxplus \top \wedge \Box A$ at a world implies the truth of $\blacksquare A$ in the model, but it does not quite give us the global modality as the converse implication does not hold in general. Rather, what we are looking for is

$$\blacksquare A := \boxplus \neg A \wedge \Box A.$$

The first conjunct tells us that every inaccessible world is a A-world and the second that every accessible world is a A-world, whence every world is a A-world.

In the presence of two modal operators, a standard Kripkean one governed by R and its complementary sufficiency one essentially governed by $\overline{R}$, the sufficiency modality allows us to talk about worlds that are R-inaccessible: the dual modality $\Diamond\!\!\!\!+ A$ says that there is an inaccessible A-world. As such, $L(\boxplus)$

lacks the usual version of invariance under generated submodels. However, while we are without the notion of generated submodel for $L(\boxplus)$, there is an analog. Informally, M' is a generated $L(\boxplus)$-submodel[17] of M if M' is a submodel of M and

$$(6.7) \qquad \text{if } a \in W' \text{ and } \neg Rab \text{ then } b \in W'.$$

Proposition 6.3.1. *Let $M' = (W', R', V')$ be a generated $L(\boxplus)$-submodel of $M = (W, R, V)$. Then for all $a \in W'$*

$$M', a \models A \text{ iff } M, a \models A.$$

Proof. By induction on the complexity of A. We do only the case for $A = \diamondplus B$, supposing the inductive hypothesis holds for all B with complexity less than A's. (Working with $\diamondplus$ over $\boxplus$ in this case is more revealing.)

Suppose $M, a \models \diamondplus B$. Then $\exists b \in W$ such that $\neg Rab$ and $M, b \models \neg B$. By (6.7), $b \in W'$, whence by IH $M', b \models \neg B$. So $M', a \models \diamondplus B$.

For the other direction suppose $M', a \models \diamondplus A$. Then $\exists b \in W'$ such that $\neg R'ab$ and $M', b \models \neg A$. Then $\neg Rab$, and by IH $M, b \models \neg A$. Whence $M, a \models \diamondplus A$. $\qquad \square$

As noted earlier, $L(\boxplus)$-formulas are not preserved under taking disjoint unions. The problem is that if $M = (W, R, V)$ and $M' = (W', R', V')$ are disjoint models and $M, a \models \boxplus B$, M' may have worlds at which B is true that are inaccessible from a, in which case $M \uplus M', a \not\models \boxplus B$. The only thing we can do to ensure that formulas involving $\boxplus$ are preserved is, after forming their disjoint union, connect all the worlds from one model to all the worlds in the others and vice versa for pairwise disjoint models. For simplicity we work in the language with a single modal operator $\boxplus$.

Definition 6.3.2. Let $\{F_i = (W_i, R_i)\}_{i \in I}$ be a family of disjoint frames for $L(\boxplus)$. Their *disjoint $L(\boxplus)$-union* $\uplus_{i \in I}^{\boxplus} F_i = (W^*, R^*)$ is formed by taking their disjoint union $\uplus_{i \in I} F_i$ and setting $R_i a_i b_j$ for all $a_i \in W_i$ and $b_j \in W_j$ for $j \neq i$. That is

1. $W^* = \uplus_{i \in I} W_i$;

2. $R^* = \uplus_{i \in I} R_i \cup \{(a_i, b_j) : a_i \in W_i, b_j \in W_j\}$ for all $i, j \in I$ s.t. $i \neq j$.

The disjoint $L(\boxplus)$-union of a family $\{M_i = (W_i, R_i, V_i)\}_{i \in I}$ of models is defined likewise by additionally setting $V^* = \uplus_{i \in I} V_i$. We obtain the following invariance results.

Theorem 6.3.3. *Let $\{F_i\}_{i \in I}$ be a family of disjoint frames for $L(\boxplus)$. Then $F_i \models A$ for each $i \in I$ iff $\bigcup_{i \in I}^{\boxplus} F_i \models A$.*

Corollary 6.3.4. *Let $\{M_i\}_{i \in I}$ be a family of disjoint models for $L(\boxplus)$. Then for each $a_i \in W_i$ and A in $L(\boxplus)$, $\uplus_{i \in I}^{\boxplus} M_i, a_i \models A$ iff $M_i, a_i \models A$.*

---

[17]The reader is referred to [BdRV02] for such notions as (generated) submodel.

Proof. We prove only the corollary. The base case is immediate and as usual we cover only the case for $A = \bigoplus B$, supposing the inductive hypothesis holds for all B with complexity less than A's.

Suppose $M_i, a_i \models \bigoplus B$. Then $\exists b \in W_i$ s.t. $M_i, b \models \neg A$ and not $R_i a_i b$. By IH, $\biguplus_{i \in I}^{\boxplus} M_i, b \models \neg A$. Since $\forall c \in W_i$, $R_i a_i c$ iff $R^* a_i c$, not $R^* a_i b$. Whence $\biguplus_{i \in I}^{\boxplus} M_i, a_i \models \bigoplus B$.

For the other direction, suppose $\biguplus_{i \in I}^{\boxplus} M_i, a_i \models \bigoplus B$. Then $\exists b \in W^*$ s.t. $\biguplus_{i \in I}^{\boxplus} M_i, b \models \neg A$ and not $R^* a_i b$. By the definition of R^*, $b_j \in W_i$ and not $R_i a_i b$. So by IH, $M_i, b \models \neg A$, whence $M_i, a_i \models \bigoplus B$ as desired. $\square$

Invariance results for $L(\boxplus)$ such as these give us an easy means of proving indefinability results just as their analogs do for Kripkean languages such as L. The following is but one example.

Example 6.3.5. Antisymmetry, i.e. if Rxy and Ryx then $x = y$, is not $L(\boxplus)$-definable.

Proof. Suppose $A \in L(\boxplus)$ defines antisymmetry. Let F_1 and F_2 be antisymmetric frames. Then $F_i \models A$ for each $i \leq 2$. By theorem 6.3.3, $\bigcup_{i \leq 2}^{\boxplus} F_i \models A$. But clearly the alternative relation R^* of $\bigcup_{i < w}^{\boxplus} F_i$ is not antisymmetric as each $a \in W_1$ is distinct from each $b \in W_2$ and yet we have $R^* ab$ and $R^* ba$ by definition. $\square$

The following theorem establishes the relationship between L- and $L(\boxplus)$-models and their theories.

Theorem 6.3.6. *Let $M = (W, R, V)$ and $M' = (W, \overline{R}, V)$ be models, with $\overline{R}$ the complement of R (relative to W), and let $Th_{L(\boxplus)}^M(a)$ be the $L(\boxplus)$-theory of $a' \in W'$ (and likewise for $a \in W$). Moreover let $t(Th_{L(\boxplus)}^{M'}(a))$ be the translation which substitutes in every $L(\boxplus)$-formula $A \in Th_{L(\boxplus)}(a^{M'})$, $\boxplus$ for $\square$ and $\bigoplus$ for $\lozenge$. Then*

$$Th_L^M(a) = t(Th_{L(\boxplus)}^{M'}(a)).$$

Informally t translates an $L(\boxplus)$-theory into an "equivalent" L-theory by replacing $L(\boxplus)$-boxes for L-boxes and likewise for diamonds. If we were to use the same symbol for the primitive modal operators of L and $L(\boxplus)$, the theorem would tell us that for every point $a \in W$, its L-theory in M is the same as its $L(\boxplus)$-theory in M'. Clearly the theorem holds translating the other way around, from L-formulas to $L(\boxplus)$-ones. As a concrete example take $\square A \to A$ which defines the reflexivity of R. The above theorem tells us that $\boxplus A \to A$ defines the *irreflexivity* of R.

Proof. It suffices to show that $M, a \models A$ iff $M', a \models A'$ where A' is exactly like A except that there is an occurrence of $\boxplus$ wherever there is an occurrence of $\square$ in A. The proof is by induction on formula complexity. We do only the case $A = \lozenge B$. (It is more revealing to work with the diamond case here.) Suppose $M, a \models \lozenge B$. Then there is a $b \in W$ such that Rab and $M, b \models B$. Then $\neg \overline{R} ab$. By IH, $M', b \models B$, whence $M', a \models \bigoplus B$.

For the other direction, suppose $M', a \models \bigoplus B$. Then there is a $b \in W$ such that $\neg \overline{R} ab$ and $M', b \models B$. Then Rab. By IH, $M, b \models B$, whence $M, a \models \lozenge B$. $\square$

We may also obtain the obvious analogs of results such as bounded morphisms (also called 'p-morphisms') and bisimulations by replacing in them R for its negation.

An axiomatization of $\Box$ and its complementary modality $\boxplus$ in the language $L(\Box\boxplus)$ is provided in [Hum83]. It consists of the K-axioms and necessitation rules for each box, along with modus ponens and the following "meta"-schema[18]:

(C) $$D(\Box A \land \boxplus\neg B) \to B(A \lor B),$$

where D and B range over (possibly empty) strings of diamonds (either $\Diamond$ or $\diamondplus$) and boxes (either $\Box$ or $\boxplus$) respectively. (Note that D here is not the difference operator of §6.3.) We call this logic **KC**, **C** for 'complementary', the completeness of which is established in [Hum83] by the less familiar method of Cresswell, rather than the usual canonical models method of Scott and Lemmon.

Theorem 6.3.7 (Soundness). *Let $\mathcal{K}$ be the class of all complementary frames $F = (W, R_1, \overline{R_1})$. Then $\mathbf{KC} \vdash A$ implies $\mathcal{K} \models A$.*

Proof. Let $F = (W, R_1, R_2)$ be a complementary frame and suppose that $M = (F, V), a \models D(\Box A \land \boxplus\neg B)$. Then there is an $R_1 \cup R_2$-chain of length n $(n < \omega)$ to a $b \in W$ s.t. $M, b \models \Box A \land \boxplus\neg B$. Thus for all c s.t. $R_1 bc$, $M, c \models A$ and for all c s.t. $R_2 bc$, $M, c \models B$ (as $R_2 = \overline{R_1}$). Suppose there is an $R_1 \cup R_2$-chain of length m from a to c. Then either $R_1 bc$ or not. If so, then $M, c \models A$ so $M, c \models A \lor B$. If not, then $R_2 bc$ so $M, c \models B$ so $M, c \models A \lor B$. Whence $M, a \models B(A \lor B)$ as m and c were arbitrary. $\square$

6.4   Final remarks

If we are not given the global modality for free, a definition of negation as absolute impossibility requires placing rather unfamiliar constraints on a typical modal language including a counterfactual operator and propositional quantification. The primary aim of the present chapter was to clarify precisely which constraints are so needed and to further investigate the definability of the global modality used in the definition of negation as impossibility in languages including a modality and its complementary.

As a final reminder, it is worth highlighting again the nearly identical negations of **SS5** of section 6.2 and **IPC**. Their similarity is not surprising given their sharing of truth conditions, but similarity to this level becomes surprising when one sees that it is accounted for by more than their mere sharing of truth conditions, a fact made evident when we consider that neither does the semantics of **SS5** impose Heredity nor does the semantics of **IPC** employ universal frames.

---

[18]Humberstone ([Hum83]) calls it a single schema which is somewhat misleading as schemas are usually understood. In the case of schema (C), one substitutes, not just formulas for schematic metavariables ranging over such formulas, but also ill-formed strings of a special sort for metavariables ranging over such strings.

Chapter 7

Final words

We looked at five issues concerning negation. Of those five there are two of an especially general character that I wish to revisit. The first of those concerns whether it is at all meaningful or legitimate to engage in revisionary debate concerning the nature of negation. I think the question has two answers, one of which was not touched upon in chapter 2. In that chapter I answered with a resounding "Yes", revisionary debate is possible, but there is a negative answer with significant appeal which runs in one of the following two ways. When the deviant rejects some classical principle or accepts some non-classical principle regarding negation, she may be seen as denying the universal validity of classical logic. That is, she may be seen as denying that classical principles are valid under absolutely *every* domain of discourse. The classicist might be happy to agree. He may not think that the domain of classical logic extends to fiction, for example. For in the world of fiction, anything goes! This kind of classicist may either reject the subject or domain neutrality of classical logic or he may demote fiction and other domains as unworthy of logical constraint. If the former, he may still hold that classical logic holds good *relative* to this or that—but not every—domain. So this breed of classicist may deny that revisionary debate is possible when the domain is "classical logic friendly", though he may nonetheless be open to debate concerning precisely on which domains classical logic holds good. This classicist is a kind of relativist who denies that revision is possible relative to some domains D but not relative to others D', perhaps because these D' are not on solid footing.

I have my sympathies with this classicist. After all, would anyone really deny that classical logic fails to preserve truth when the domain of discourse is fiction or what is believed, doubted, etc. by Frank to be true? Surely not. Yet it is not clear that these domains are somehow illegitimate as regards logic so we should give them their fair due. But to do so is to be open to revising our theories of what constitutes valid inference, and in particular when those inferences concern negation, in a doxastic, fictional or other setting. Whether such domains are taken to be legitimate is a deep question regarding the *demarcation* of logic—Is the "logic" of fiction logic?—and it is intimately intertwined with the question of whether revisionary debate is possible. I think it is possible, but on account of enough shared meaning between the classicist and deviant's respective uses of the relevant logical vocabulary such as 'not'.

The second issue of an especially broad character that we addressed con-

cerns truthmakers for negative truths. On one extreme we have truthmaker maximalism. The principle strikes me as characteristic of a naïve theory of truthmaking, one that is bound to get us into hot water if we do no heavy tinkering in other areas such as logic. This is a familiar affair with naïve theories; just consider naïve set theory, naïve theory of properties and predicates, naïve truth theory and so forth. On the other extreme we have atomism. This view strikes me as unnecessarily restrictive, analogous to solving the liar paradox by banishing sentential self-reference. (At the level of propositions I think this is actually the right thing to say. But it is not at the level of sentences.) So there is a forceful intuition that it must be some theory in between these extremes that is the correct one. The hard part is giving a precise account of this intermediate theory. I hope to have gone all the way toward this end in chapter 5 by arguing for a version of maximalism restricted to positive truths and by further providing the essential ingredient that is an account of positivity.

Bibliography

[AB63]   A. R. Anderson and N. D. Belnap JR. First degree entailments. *Mathematische Annalen*, 149(4):302–319, 1963.

[Arm89]   D. M. Armstrong. *Universals: An Opinionated Introduction.* Focus Series. Westview Press, 1989.

[Arm04]   D. M. Armstrong. *Truth and Truthmakers.* Cambridge University Press, 2004.

[Aye52]   A. J. Ayer. Negation. *The Journal of Philosophy*, 49(26):797–815, 1952.

[Bax88]   Donald L. M. Baxter. Identity in the loose and popular sense. *Mind*, 97(388):575–582, 1988.

[BdRV02]   Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic.* Cambridge University Press, 2002.

[BE87]   Jon Barwise and John Etchemendy. *The Liar: An Essay on Truth and Circularity.* Oxford University Press, 1987.

[Bèz02]   J. Y. Bèziau. S5 is a paraconsistent logic and so is first-order classical logic. *Logical Studies*, 9:1–9, 2002.

[Bla64]   Max Black. *A Companion to Wittgenstein's Tractatus.* Cambridge University Press, 1964.

[Bog92]   James Bogan. Change and contrariety in Aristotle. *Phronesis*, 37(2):1–21, 1992.

[Bra93]   David Braun. Empty names. *Noûs*, 27(4):449–469, December 1993.

[Cam08]   Ross Cameron. How to be a truthmaker maximalist. *Noûs*, 42(3):410–421, 2008.

[CC00]   Roy T. Cook and Jon Cogburn. What negation is not: Intuitionism and $0 = 1$. *Analysis and Metaphysics*, 60(1):5–12, 2000.

[CCCM01]   W. Carnielli, C. Caleiro, M. E. Coniglio, and J. Marcos. Suszko's thesis and dyadic semantics. 2001. unpublished ms.

[Che75]   Brian Chellas. Basic conditional logic. *Journal of Philosophical Logic*, 4:133–153, 1975.

[Cop79]   B. J. Copeland. On when a semantics is not a semantics: Some reasons for disliking the routley-meyer semantics for relevance logic. *Journal of Philosophical Logic*, 8(November):399–413, 1979.

[Cop83]   B. J. Copeland. Pure semantics and applied semantics. *Topoi*, 2(2):197–204, December 1983.

[Cop86]   B. J. Copeland. What is a semantics for classical negation? *Mind*, 95(380):478–490, October 1986.

[Cre88]   M. J. Cresswell. Necessity and contingency. *Studia Logica*, 47:145–149, 1988.

[DH01]   Michael J. Dunn and Gary Hardegree. *Algebraic methods in philosophical logic*. Oxford University Press, 2001.

[Dos86]   Kosta Došen. Negation as a modal operator. *Reports on Mathematical Logic*, 20:15–27, 1986.

[Dos99]   Kosta Došen. Negation in the light of modal logic. In *What is Negation?*, Applied Logic Series, pages 77–86. Kluwer Academic Publishers, 1999.

[Dod02]   Julian Dodd. Is truth supervenient on being? *Proceedings of the Aristotelian Society*, 102:69–85, 2002.

[Dod07]   Julian Dodd. Negative truths and truthmaker principles. *Synthese*, 156(2):383–401, 2007.

[dR92]   Maarten de Rijke. The modal logic of inequality. *The Journal of Symbolic Logic*, 57(2):566–584, June 1992.

[DS06]   David DeVidi and Graham Solomon. Empirical negation in intuitionisitic logic. In David DeVidi and Tim Kenyon, editors, *A Logical Approach to Philosophy*, pages 151–168. Springer, Dordrecht, The Netherlands, 2006.

[Dum77]   Michael Dummett. *Elements of Intuitionism*. Oxford University Press, USA, second edition, 1977.

[Dum91]   Michael Dummett. *The Logical Basis of Metaphysics*. Harvard University Press, 1991.

[Dum96]   Michael Dummett. *The seas of language*. Oxford university press, 1996.

[Dum09]   Michael Dummett. Fitch's paradox of knowability. In *New Essays on the Knowability Paradox*, pages 51–52. Oxford University Press, 2009.

[Dun93]   Michael M. Dunn. Star and perp: Two treatments of negation. *Philosophical Perspectives*, 7:331–357, 1993.

[Dun05]   J. Michael Dunn. Negation in the context of gaggle theory. *Studia Logica*, 80:235–264, 2005.

[Dur98]   Michael Durrant. Plato's Quinean beard: Did Plato ever grow it? *Philosophy*, 73:113–121, 1998.

[Fie08]   Hartry Field. *Saving truth from paradox*. Oxford University Press, USA, 2008.

[Fin70]   Kit Fine. Propositional quantification in modal logic. *Theoria*, 36:336–346, 1970.

[Fin94]   Kit Fine. Essence and modality. *Philosophical Perspectives*, 8:1–16, 1994.

[Fre92]   Gottlob Frege. On concept and object. *Vierteljahrsschrift fur wissenschaftliche Philosophie*, 16:192–205, 1892. Translation by P. T. Geach, revised by Max Black.

[Fre36]   Hans Freudenthal. Zur intuitionistischen deutung logischer formeln. *Compositio Mathematica*, 4(1):112–116, 1936.

[Fre48]   Gottlob Frege. Sense and reference. *Philosophical Review*, 57(3):209–230, May 1948. originally dated 1892.

[Fre60]   Gottlob Frege. Negation. In Peter Geach and Max Black, editors, *Translations from the Philosophical Writings of Gottlob Frege*. Basil Blackwell, 2nd edition, 1960.

[Fre63]   Frederic B. Fitch. A logical analysis of some value concepts. *The Journal of Symbolic Logic*, 28(2):135–142, June 1963.

[Gar01]   James W. Garson. Quantification in modal logic. In D. M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic (2nd ed.), Vol. 3*, pages 267–323. Kluwer Academic Publishers, Netherlands, 2001.

[Gea69]   Peter Geach. Contradictories and contraries. *Analysis*, 29(6):187–190, 1969.

[GH99]   Dov Gabbay and Anthony Hunter. Negation and contradiction. In Dov Gabbay and Heinrich Wansing, editors, *What is Negation?*, pages 89–100. Kluwer Academic Publishers, 1999.

[Gol74]   R. I. Goldblatt. Semantic analysis of orthologic. *Journal of Philosophical Logic*, 3(1):19–35, March 1974.

[Gor90]   Valentin Goranko. Modal definability in enriched languages. *Notre Dame Journal of Formal Logic*, 31(1):81–105, Winter 1990.

[GP92]   Valentin Goranko and Solomon Passy. Using the universal modality: Gains and questions. *Journal of Logic and Computation*, 2(1):5–30, 1992.

[GPT87]   G. Gargov, S. Passy, and T. Tinchev. Modal environment for boolean speculations. In D. Scordev, editor, *Mathematical Logic and Its Applications*, chapter pp. 253-263. Plenum Press, New York, 1987.

[Gri46]   G. F. C. Griss. Negationless intuitionistic mathematics. *I, Koninklijke Nederlandsche Akademie van Wetenschappen, Proceedings of the Section of Sciences*, 49:1127–1133, 1946.

[Gur77]   Yuri Gurevich. Intuitionistic logic with strong negation. *Studia Logica*, 36(1/2):49–59, 1977.

[Haz95]   A. P. Hazen. Is even minimal negation constructive? *Analysis*, 55(2):105–107, April 1995.

[HBW08]   Hans van Ditmarsch, Barteld Kooi, and Wiebe van der Hoek. *Dynamic epistemic logic*. Springer, 2008.

[Hei43]   F. H. Heinemann. The meaning of negation. *Proceedings of the Aristotelian Society*, 44:127–152, 1943.

[Hey55]   A. Heyting. G. F. C. Griss and his negationless mathematics. *Synthese*, 9(1):91–96, 1955.

[Hey71]   A. Heyting. *Intuitionism: An Introduction*. North-Holland Publishing Co., Amsterdam, third edition, 1971.

[Hod84]   Harold T. Hodes. Some theorems on the expressive limitations of modal languages. *Journal of Philosophical Logic*, 13:13–26, 1984.

[Hor01]   Laurence R. Horn. *A Natural History of Negation*. CSLI Publications, 2001.

[Hos90]   Keith G. Hossack. A problem about the meaning of intuitionist negation. *Mind*, 99(394):207–219, April 1990.

[Hum83]   Lloyd Humberstone. Inaccessible worlds. *Notre Dame Journal of Formal Logic*, 24(3):346–352, July 1983.

[Hum87]   Lloyd Humberstone. The modal logic of 'all and only'. *Notre Dame Journal of Formal Logic*, 28(2):177–188, 1987.

[Hum95]   Lloyd Humberstone. The logic of non-contingency. *Notre Dame Journal of Formal Logic*, 36:214–229, 1995.

[Hum00]   Lloyd Humberstone. The revival of rejection negation. *Journal of Philosophical Logic*, 29:331–381, 2000.

[Hum03]   Lloyd Humberstone. Note on contraries and subcontraries. *Noûs*, 37(4):690–705, 2003.

[Jag10]   Mark Jago. Setting the facts straight. *Journal of Philosophical Logic*, online first:1–22, June 2010.

[Kan03]   Akihiro Kanamori. The empty set, the singleton, and the ordered pair. *The Bulletin of Symbolic Logic*, 9(3):273–298, September 2003.

[Kap78]   David Kaplan. On the logic of demonstratives. *Journal of Philosophical Logic*, 8:81–98, 1978.

[Kin95]   Jeffrey C. King. Structured propositions and complex predicates. *Noûs*, 29(4):516–535, 1995.

[Kle45]   S. C. Kleene. On the interpretation of intuitionistic number theory. *The Journal of Symbolic Logic*, 10(4):109–124, December 1945.

[Kne56]   William Kneale. *The Province of Logic*, pages 237–261. The University Press Aberdeen, 1956.

[Kri75]   Saul Kripke. Outline of a theory of truth. *The Journal of Philosophy*, 72(19):690–716, 1975.

[KT97]   Michael Kaminski and Michael Tiomkin. The expressive power of second-order propositional modal logic. *Notre Dame Journal of Formal Logic*, 37(1):35–43, Winter 1997.

[Kuh95]   Steven T. Kuhn. Minimal non-contingency logic. *Notre Dame Journal of Formal Logic*, 36(2):230–234, 1995.

[Kuk06]   Boris Kukso. The reality of absences. *Australasian Journal of Philosophy*, 84(1):21–37, March 2006.

[Leb76]   Hughes Leblanc. *Truth-value Semantics*. North-Holland Publishing Co., 1976.

[Len96]   Wolfgang Lenzen. Necessary conditions for negation operators. In Heinrich Wansing, editor, *Negation: A Notion in Focus*, Perspective in Analytical Philosophy, pages 37–58. Walter de Gruyter, 1996.

[Lew73]   David Lewis. *Counterfactuals*. Harvard University Press, Cambridge, Massachusetts, 1973.

[Lew76]   David Lewis. Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, 85(3):297–315, July 1976.

[Lew86]   David Lewis. *On the Plurality of Worlds*. Blackwell Publishing, 1986.

[Lew91]   David Lewis. *Parts of Classes*. Blackwell, 1991.

[Lew01]   David Lewis. Truthmaking and difference-making. *Noûs*, 35(4):602–615, 2001.

[Lew03]   David Lewis. *Things Qua Truthmakers*, pages 25–38. Routledge, London and New York, 2003.

[Mar84]   Per Martin-Löf. *Intuitionistic Type Theory*. Bibliopolis, 1984.

[McC67]   Storrs McCall. Contrariety. *Notre Dame Journal of Formal Logic*, 8(1):121–132, April 1967.

[Mel03]   D. H. Mellor. Real metaphysics: Replies. In *Real Metaphysics: Essays in Honour of D. H. Mellor*, pages 212–237. Routledge, 2003.

[Mol00]   Geroge Molnar. Truthmakers for negative truths. *Australasian Journal of Philosophy*, 78(1):72–86, 2000.

[Mor76]   Chris Mortensen. A sequence of normal modal systems with non-contingency bases. *Logique et Analyse*, 19:341–344, 1976.

[Mor89]   Chris Mortensen. Anything is possible. *Erkenntnis*, 30:319–337, 1989.

[MR66]   H. Montgomery and R. Routley. Contingency and non-contingency bases for normal modal logics. *Logique et Analyse*, 9:318–328, 1966.

[MR73]   Robert K. Meyer and Richard Routley. Classical relevant logics i. *Studia Logica*, 32(1):51–66, December 1973.

[MR74]   Robert K. Meyer and Richard Routley. Classical relevant logics ii. *Studia Logica*, 33(2):183–194, June 1974.

[MSS84]   Kevin Mulligan, Peter Simons, and Barry Smith. Truth-makers. *Philosophy and Phenomenological Research*, 44(3):287–321, 1984.

[MT48]   J. C. C. McKinsey and Alfred Tarski. Some theorems about the sentential calculi of Lewis and Heyting. *The Journal of Symbolic Logic*, 13(1):1–15, March 1948.

[Mum05]   Stephen Mumford. The true and the false. *Australasian Journal of Philosophy*, 83(2):263–269, 2005.

[Nel49]   David Nelson. Constructible falsity. *The Journal of Symbolic Logic*, 14(1):16–26, March 1949.

[Par84]   Terence Parsons. Assertion, denial and the liar paradox. *Journal of Philosophical Logic*, 13:137–152, 1984.

[Par99]   Josh Parsons. There is no 'truthmaker' argument against nominalism. *Australasian Journal of Philosophy*, 77(3):325–334, September 1999.

[Pri79]   Graham Priest. The logic of paradox. *Journal of Philosophical Logic*, 8(1):219–241, January 1979.

[Pri93]   Graham Priest. Can contradictions be true? *Proceedings of the Aristotelian Society, Supplementary Volumes*, 67:34–54, 1993.

[Pri99]   Graham Priest. What not? a defence of dialetheic theory of negation. In Heinrich Wansing, editor, *What is Negation?*, pages 101–120. Kluwer Academic Publishers, 1999.

[Pri06]   Graham Priest. *Doubt Truth to be a Liar*. Oxford University Press, USA, 2006.

[Qui51]   W. V. Quine. Two dogmas of empiricsm. *The Philosophical Review*, 60(1):20–43, 1951.

[Qui70]   W. V. Quine. *Philosophy of Logic*. Prentice Hall Inc., 1970.

[Rea94]   Stephen Read. *Thinking About Logic*. Oxford University Press, 1994.

[Rea00]   Stephen Read. Harmony and autonomy in classical logic. *Journal of Philosophical Logic*, 29:123–154, 2000.

[Rea05]   Stephen Read. Unity of the factthe 80. *The Royal Institute of Philosophy*, 80:317–342, 2005.

[Rea08]   Stephen Read. Harmony and modality. In Cédric Dégremont, Laurent Keiff, and Helge Rückert, editors, *Dialogues, Logics and Other Strange Things: Essays in Honour of Shahid Rahman*, pages 285–304. College Publications, 2008.

[Rea10]   Stephen Read. Necessary truth and proof. *forthcoming*, pages 1–21, 2010.

[Res96]   Greg Restall. Truthmakers, entailment and necessity. *Australasian Journal of Philosophy*, 74(2):331–340, 1996.

[Res99]   Greg Restall. *Negation in Relevant Logics: How I Stopped Worrying and Learned to Love the Routley Star*. Applied Logic Series. Kluwer Academic Publishers, 1999.

[Res05]   Greg Restall. Multiple conclusions. In Petr Hájek, Luis Valdes-Villanuiva, and Dag Westerståhl, editors, *Logic, Methodology and Philosophy of Science*. College Publications, 2005.

[Res08]   Greg Restall. Modal models for bradwardine's theory of truth. *The Review of Symbolic Logic*, 1(Special Issue 02):225–240, 2008.

[RM77]   Richard Routley and Robert K. Meyer. Extensional reduction i. *Monist*, 60(3):355–369, 1977.

[Ros72]   Jay F. Rosenberg. Russell on negative facts. *Noûs*, 6(1):27–40, March 1972.

[Rou75]   Richard Routley. Universal semantics? *Journal of Philosophical Logic*, 4:327–356, 1975.

[Rou84]   Richard Routley. The american plan completed: Alternative classical style semantics, without stars, for relevant and paraconsistent logics. *Studia Logica*, 43:131–158, 1984.

[Rum97]   Ian Rumfitt. The categoricity problem and truth-value gaps. *Analysis*, 57(4):223–235, October 1997.

[Rum00]   I. Rumfitt. 'yes and no'. *Mind*, 109(436):781–823, October 2000.

[Rus40]   Bertrand Russell. *An Enquiry into Meaning and Truth*. 1940.

[Rus10]   Bertrand Russell. *On the Philosophy of Logical Atomism*. Routledge Classics, 2010. originally dated 1918.

[Sal95]   Nathan Salmon. Being of two minds: Belief with doubt. *Noûs*, 29(1):1–20, 1995.

[San68]   David H. Sanford. Contraries and subcontraries. *Noûs*, 2(1):95–96, 1968.

[Seu10]   Pieter A. M. Seuren. *The Logic of Language: Language from Within*, volume II. Oxford University Press, 2010.

[Sid96]   Ted Sider. Naturalness and arbitrariness. *Philosophical Studies*, 81:283–301, 1996.

[Sim05]   Peter Simons. Negatives, numbers, and necessity some worries about armstrong's version of truthmaking. *Australasian Journal of Philosophy*, 83(2):253–261, 2005.

[Sla95]   B. H. Slater. Paraconsistent logics? *Journal of Philosophical Logic*, 24(4):451–454, August 1995.

[Smi93]   Timothy Smiley. Can contradictions be true? *Proceedings of the Aristotelian Society, Supplementary Volumes*, 67:17–33, 1993.

[Smi96]   Timothy Smiley. Rejection. *Analysis*, 56(1):1–9, January 1996.

[Smi99]   Barry Smith. Truthmaker realism. *Australasian Journal of Philosophy*, 77(3):274–291, 1999.

[ST70]   Robert C. Stalnaker and Richmond H. Thomason. A semantic analysis of conditional logic. *Theoria*, 36(1):23–42, 1970.

[Sta68]   Robert C. Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Blackwell Publishers, Oxford, 1968.

[Sus76]   Roman Suszko. The Fregean axiom and Polish mathematical logic in the 1920s. *Studia Logica*, 36(4):377–380, 1976.

[SW10]   Yaroslov Shramko and Heinrich Wansing. Truth values. *Stanford Encyclopedia of Philosophy*, 2010.

[Tal09]   Jonathan Tallant. Presentism and truth-making. *Erkenntnis*, 71:407–416, 2009.

[Ten97]   Neil Tennant. *The Taming of the True*. Oxford University Press, New York, 1997.

[Tho69]   R. H. Thomason. A semantical study of constructible falsity. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 15:247–257, 1969.

[Tho76]   S. K. Thomason. Possible worlds and many truth values. *Studia Logica*, 37(2):195–204, 1976.

[Tom06]   Paul Tomassi. Truth, warrant and superassertibility. *Synthese*, 148:31–56, 2006.

[Tsu98]   Marcelo Tsuji. Many-valued logics and suszko's thesis revisited. *Studia Logica*, 60(2):299–309, 1998.

[vBRvO06] Johan van Benthem, Olivier Roy, and Sieuwert van Otterloo. Preference logic, conditionals, and solution concepts in games. In *Modality matters: twenty-five essays in honor of Krister Segerberg, Uppsala Philosophical Studies 53*. Uppsala, 2006.

[VC09] Achille C. Varzi and Roberto Casati. Holes. *Stanford Encyclopedia of Philosophy*, 2009.

[vD08] Dirk van Dalen. *Logic and Structure*. Springer, second edition, 2008.

[Vic00a] Victor N. Krivtsov. A negationless interpretation of intuitionistic theories I. *Studia Logica*, 64:323–344, 2000.

[Vic00b] Victor N. Krivtsov. A negationless interpretation of intuitionistic theories II. *Studia Logica*, 65:155–179, 2000.

[W. 61] W. V. Quine. *From a Logical Point of View: Logico-Philosophical Essays*. Harper and Row, New York, 2nd edition, 1961.

[Wan96] Heinrich Wansing, editor. *Negation: A Notion in Focus*. Perspectives in Analytical Philosophy. Walter de Gruyter, Berlin, Germany, 1996.

[Wan02] Heinrich Wansing. Diamonds are a philosopher's best friend. *Journal of Philosophical Logic*, 31:591–612, 2002.

[WG99] Heinrich Wansing and Dov M. Gabbay, editors. *What is Negation?* Applied Logic Series. Springer, 1999.

[Wil94] Timothy Williamson. Never say never. *Topoi*, 13(2):135–145, 1994.

[Wit21] Ludwig Wittegenstein. *Tractatus Logico-Philosophicus*. Routledge, Pears & McGuinness edition, 1961 (1921).