

The Metaphysical Commitments of Logic

Thomas Nicolaas Pieter Albert Brouwer

University of Leeds

School of Philosophy, Religion and History of Science

**Submitted in accordance with the requirements for the
degree of Doctor of Philosophy**

June 2013

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2013 The University of Leeds and Thomas Nicolaas Pieter Albert Brouwer

The right of Thomas Nicolaas Albert Brouwer to be identified as the Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act of 1988.

Acknowledgements

First and foremost, I want to thank my supervisors, Robbie Williams, Jason Turner and Robin Le Poidevin. Without their guidance, this thesis would be nothing like it is, though all of its shortcomings are of course my fault. Their commitment, interest and insight have made my project a pleasure to complete. Apart from my supervisors, I am also indebted to the members of the Centre for Metaphysics and Mind for philosophical conversation both on topics related to my thesis and other subjects. Special mention goes to Ross Cameron, Andy McGonigal and Paolo Santorio in this regard. Appreciation goes to Aaron Meskin for his persistent efforts to get me to give a talk on aesthetics, which I finally did. I want to especially thank John Divers for his support.

The postgraduate community in Leeds is incomparable. I want to thank all of my fellow postgrads for helping to make Leeds a great place to do a PhD, and my fellow philosophers in particular for endless discussion on all philosophical topics. For listening to me trying out arguments and views and helping me out on many occasions, I want to particularly thank Jon Banks, Jordan Bartol, Callum Sida-Murray and Carl Warom. Fond thanks also go to Sarah 'plain old' Adams, Dani Adams, Richard Caves, Dave Cornell, Fran Fairbairn, Mirja Holst, Wouter Kalf, Alice Kay, Kerry McKenzie, Henry Merivale, Robb Muckle, Alex Oldemeier, Andy Peet, Will Perry, Rob Pezet and Paul Ramshaw. Most of all, I want to thank Mike Bench-Capon, who is not only clever himself but the cause of cleverness in others. And special mention goes to Efram Sera-Shriar, who did much to make me feel at home when I first arrived in Leeds.

I also want to express my gratitude to the administrative staff at the School, for always being ready to help out a confused postgrad, and always there for a friendly chat. In particular I want to thank Sarah Cason for always cheering me up.

In a category of her own, I want to thank Akiko Frischhut for good philosophy, good food and good company, and for making Geneva one more home away from

home.

Finally, I want to thank my family back home in the Netherlands, who have provided so much moral support and continue to do so. I could not have done it without them. I hope they will forgive me for being so far away for such a long time.

Abstract

This thesis is about the metaphysics of logic. I argue against a view I refer to as ‘logical realism’. This is the view that the logical constants represent a particular kind of metaphysical structure, which I dub ‘logico-metaphysical structure’. I argue instead for a more metaphysically lightweight view of logic which I dub ‘logical expressivism’.

In the first part of this thesis (Chapters I and II) I argue against a number of arguments that Theodore Sider has given for logical realism. In Chapter I, I present an argument of his to the effect that logico-metaphysical structure provides the only good explanation of the semantic determinacy of the logical constants. I argue that other explanations are possible. In Chapter II, I present another argument of his to the effect that logico-metaphysical structure is something that comes along with ontological realism: the view that there is a non-language-relative fact of the matter about what exists. I argue that the connection between logical and ontological realism is not as close as Sider makes it out to be.

In the second part of this thesis (Chapters III – V) I set out a positive view of the logical constants that can explain both why their meanings are semantically determinate, and why they form part of our vocabulary. On that view, the primary bearers of logical structure are propositional attitudes, and the logical constants are in our language as vehicles for the expression of logically complex propositional attitudes. In Chapter III, I set out an expressivist theory of propositional logic. In Chapter IV, I use this theory to explain how the logical connectives end up having determinate meanings. In Chapter V, I extend the expressivist theory to predicate logic.

Table of Contents

Introduction	1
I. The Metaphysics of Logic	1
II. The Metaphysical Commitments of Logic	2
III. Logical Realism	4
IV. Sider's Arguments for Logical Realism	6
V. The 'Quinean' Argument from Ideological Commitment	7
VI. The Structure of the Thesis	11
Chapter I: The Argument from Semantic Determinacy	15
0. Introduction	15
I. The Indeterminacy of Meaning	17
I.I. The Kripkenstein Problem	18
II. Interpretation and Meaning	24
II.I. Lewis's Eligibility Response	25
III. A Kripkenstein Argument for Logical Constants	30
III.I. Kripkenstein Connectives	31
III.II. Kripkenstein Quantifiers	34
VI. Eligibility Responses	36
V. Conclusion	39
Chapter II: Quantifiers and Metaontology	41
0. Introduction	41
I. Quantifiers and the Realism-Deflationism Debate	43

II. The Meaning of “ \exists ”	44
III. The Structure of the Arguments	47
IV. Hirsch's Argument: a Closer Look	50
V. Eligibility and Apriority	54
VI. The Sider-Hirsch Dialectic	60
VI.I. What Ought we Disagree About?	61
VI.II. Verbal Debates about Real Facts	63
VII. Semantics and Metasemantics	64
VII.I. Local and Global Quantifier Meanings	65
VII.II. Domain Assignments	69
VIII. Conclusion	70
IX. Appendix: Unrestricted Quantification	71
Chapter III: Logical Expressivism	77
0. Introduction	78
0.I. Some Preliminaries and Caveats	81
I. Propositional Attitudes, Simple and Complex	84
I.I. Rejection	84
I.I.II. A Spectrum of Acceptance and Rejection	87
I.II. Complex Propositional Attitudes	89
I.III. Why Complex Attitudes Are Propositional Attitudes	91
I.IV. Generating Greater Complexity	93
II. Sequentese	95
II.I. The Symbol “ \blacktriangleright ”	96
II.II. Some Connectives Replaced	97
II.III. Embedded Sequents	100
III. Pinning Down Sequentese	102
III.I. Syntax of Sequentese	103
III.II. Translating Sequentese	104

III.II.I. Translating Sequentese: Sequentese to English	105
III.II.II. Translating Sequentese: from PL to Sequentese	108
III.II.III. Translating Sequentese: from Sequentese to PL	111
III.II.IV. Translating Sequentese: Recoverability	113
IV. Reasoning in Sequentese	115
IV.I. Motivating the Inference Rules	118
IV.II. Motivating the Structural Rules	124
IV.III. Embedding and Extraction Rules for Sequentese	129
IV.IV. Is Sequentese Classical?	132
IV.IV.I. Classical Derivability Implies Sequentese Derivability	133
IV.IV.II. Sequentese Derivability Implies Classical Derivability	136
V. Conclusion	141
Chapter IV: Radical indeterminacy revisited	145
0. Introduction	145
I. Chapter I Revisited	147
II. Metasemantics and Indeterminacy	149
II.I. A Lewisian Approach	149
II.II. Radical Interpretation	150
III. Lewis and Naturalness	155
IV. The Metasemantics of the Logical Connectives	158
IV.I. Psychological Assumptions	159
IV.II. Assumptions about Language	162
IV.III. Pinning Down the Constants	163
V. Reversing the Polarity on the Reference Magnets	165
VI. Conclusion	169

Chapter V: Expressivism for Quantifiers	170
0. Introduction	170
I. Strategies	172
I.I. Tractarian Quantification	173
I.II. Ramsey-Mill Quantification	175
I.II.I. Predication	176
I.II.II. Quantification	179
I.III. Correctness	181
I.IV. On the Way to Predicate Sequentese	183
II. Sequentese with Predicates: Syntax	185
II.I. Predication	185
II.II. Open and Closed Sequents	186
III. Sequentese with Predicates: Translation Manual	188
IV.I Sequentese with Predicates: Inference Rules	190
IV.II. Justifying the Inference Rules	197
V. Sequentese with Predicates is Standard Predicate Logic	198
Conclusion	201
I. Explanatory Tasks	201
II. Logical Expressivism: Challenges and Applications	203
II.I. Challenges	204
II.II. Applications	205
Bibliography	211

Introduction

I. The Metaphysics of Logic	1
II. The Metaphysical Commitments of Logic	2
III. Logical Realism	4
IV. Sider's Arguments for Logical Realism	6
V. The 'Quinean' Argument from Ideological Commitment	7
VI. The Structure of the Thesis	11

I. The Metaphysics of Logic

This thesis is about the metaphysics of logic. It investigates what we can conclude from the fact that we have the logical constants in our language, that we use them all the time, and that they are so useful to us. Does it tell us anything about the world we live in? Does the world itself have some sort of logical structure to it? Are our theories of logic in some sense theories about the world?

I defend a negative answer to these questions. Our logical terms do not reflect a special feature of reality. Logic is, in the end, all about us. We, of course, are part of the world. But there is also a sense in which we're not part of it. We can stand apart and look at the world and consider it as a thing separate from us. Insofar as we are creatures that can do that, logic belongs with us and not with the world, in the grand scheme of things.

Not everyone agrees with that sort of picture. Ted Sider, in recent years, has been defending a view on which there is most definitely logical structure to the world,¹ a view which I will be referring to as *logical realism*. Much of this thesis is concerned with countering Sider's arguments for logical realism. It will turn out that the only way of properly undermine Sider's arguments is to come up with a specific counter-

1. Sider (2009a) and (2011).

proposal, one that is metaphysically lightweight but has as much (or more) explanatory power as Sider's view does.

In this introduction, I will do a bit of ground-clearing and setting up. In section II, I discuss a few different questions one might find oneself asking under the heading of 'the metaphysical commitments of logic'. I discuss these to set them aside. In section III, I describe the nature of the view that I am concerned to argue against. In section IV I discuss Ted Sider's various arguments for the existence of logical structure, and I devote section V to the discussion of one such argument, one that, unlike the others, will not get a chapter of its own in the thesis. Finally, in section VI, I go over the structure of the thesis.

II. The Metaphysical Commitments of Logic

Before saying exactly what the question is that this thesis will try to answer, it may be helpful to distinguish a few different ways to understand the phrase "the metaphysical commitments of logic", so that I can set these aside.

First, one might understand it as being about the *ontological* commitments of logic: the things that we must accept the existence of, given our use of logic. One might think, for instance, that logic commits us to the existence of sets, given that the standard model-theoretic semantics for the logical vocabulary makes heavy use of set-theoretic vocabulary. Or one might think that we are not committed to these things, perhaps because the set-theoretic talk can be avoided, or perhaps because one believes that set-theoretic talk never commits anyone to anything.

Although I think this is a perfectly acceptable way to understand the phrase "metaphysical commitments of logic", this is not the debate I'm going to be talking about. I'm going to leave it entirely open whether logic has ontological commitments. In setting out my logical expressivism, I will here and there use set-theoretic notions, though it might be possible to do without those, if need be. I am not going to assume that my use of these notions saddles me with a commitment to sets, nor will I assume the opposite.

Second, one might wonder about the commitments that our talk of *logical consequence* – talk about what follows from what – imposes on us. Ought we to believe, for instance, that there is a real, theory-independent fact of the matter about what follows from what, a fact of the matter that we're trying to capture with our logical theories? Or ought we to think that this is merely a theory-relative matter: that there is a classical consequence relation, and a relevant one, and an intuitionist one, but that there is no such thing as 'the' consequence relation?²

This is also not a debate in which I will involve myself directly, though unlike the issue of ontological commitment, I will end up saying a thing or two about it in the course of the thesis. It will turn out that according to logical expressivism, there is a theory-independent fact of the matter about what follows from what. A further question then arises about the facts in virtue of which the one true consequence relation is what it is. According to logical expressivism, the facts in question are facts about theoretical rationality.

Third, one might wonder about the 'axiological' commitments of logic. Does logic give rise to norms on thought? If so, does it do so essentially? Are facts about what follows from what ultimately just facts about how we should think and what we should believe? Or does it do so merely accidentally, in the sense that physics imposes some norms on thought because we have a general reason to adjust our thinking to the way the world works, physically?³

It may be debated whether this question falls within the metaphysics of logic. I am inclined to count it as such. But whether it is or not, it is also not a question I'm directly concerned with in this thesis. I will, however, say something about it along the way. According to logical expressivism, logic is indeed essentially normative, though the normativity is ultimately the familiar normativity of rationality. The

2. This is a question that matters a great deal for the debate between *logical monists* and *logical pluralists*. Traditionally, it has been the philosophers who believed that there are theory-independent facts of the matter about what follows from what that were the monists – philosophers that believe that there is 'one true logic'. Recently, people have tried to combine such a non-deflationism about logical consequence with pluralism: they have argued that there can be several theory-independent consequence relations. See Beall and Restall (2005) for a non-deflationist pluralist view.

3. And if logic is essentially normative, what should we think about the status of these norms? Ought we to be naturalists, non-naturalists, or perhaps non-cognitivists about them? Resnik (1999) for instance defends an expressivist theory of the normativity of logic.

metaphysical status of that sort of normativity is something I leave to others to discuss.

I've probably not exhausted the options for interpreting the phrase "the metaphysical commitments of logic". But these are a few prominent options, and hopefully, by setting them aside, I have prevented some confusions from arising.

III. Logical Realism

In recent years, Ted Sider has been developing something of a grand unifying theory of metaphysics. In this theory, the notion of *structure* looms large. Structure, for Sider, is a technical notion, but also a theoretical primitive. He claims it can be understood on the basis of the theoretical roles that it plays.⁴

The primary theoretical role that structure has for Sider is to distinguish topics about which there are non-language relative, non-theory-relative, non-interest-relative facts of the matter from topics about which there are no such facts. If there is moral structure in the world, then there is a real fact of the matter about what's right and what's wrong. If there is modal structure, then there is a real fact of the matter about what's possible and what's impossible, what's contingent and what's necessary. In a nutshell we might say that we should believe in structure for those topics about which we are realists, at least in a certain sense of 'realist'.

Sider likes to associate bits of structure ('joints in nature', in the Platonic metaphor) with bits of vocabulary. If there is modal structure, this means that the modal operators pick out bits of the structure of reality. If there is mereological structure, then terms like 'is a part of' pick out bits of the structure of reality. Those terms then 'carve at the joints'. This association of terms with bits of structure allows Sider to formulate claims about metaphysical fundamentality in terms of the notion of structure. For if we take only the terms that have bits of structure corresponding to them ('structural terms') and if we were to describe the world using those terms, then we'd get a description of the world as it *fundamentally* is, in

4. The following is entirely based on Sider (2012), chapters 1-3.

itself, free from the distortions and baroque excrescences that our language and interests impose on it.

This leads to a second major theoretical role that structure plays for Sider. Structure is *reference magnetic*. There are various ways of carving up the contents of the world. For instance, we can divide up the objects according to whether they're green or not green, or we could carve them up according to whether they're grue or not grue. Some of these ways of dividing things up accord with the structure of reality, others don't. The green things genuinely 'go together' in a way that the grue things don't. There's objectively something better about dividing things up according to the structure of reality. It gets things more right.

Now if the terms of our language were to pick out bits of the structure of reality, then by applying those terms correctly, we would in effect divide things up according to the way reality itself divides them up. That would be a good thing. Sider considers this is a reason to think that our terms do have a tendency to pick out bits of the structure of reality. When the facts about how we use a certain colour term don't quite determine whether that term picks out the property of being green or the property of being grue, then it will in fact pick out the property of being green, since that would be the more joint-carving meaning of the two. Greenness, when compared to grueness, is a 'reference magnet'.

Sider thinks that the notion of structure plays other theoretical roles as well, to do with induction and confirmation, but I won't go into those. The important thing for our purposes is that Sider believes that the logical constants are structural. When we use the connectives and quantifiers, we thereby carve at some joints. Sider refers to the structure picked out by the logical constants as 'logical structure', but since that term has some scope for generating confusion in contexts where we want to talk about the logical structure of some sentence or proposition, I will instead be referring to it as 'logico-metaphysical structure'. It's not pretty, but it'll be clearer.

I will discuss his reasons for believing in logico-metaphysical structure in the next two sections, and at length in Chapters I and II. But before I go into that, there is one important thing to note: Sider explicitly restricts his commitment to logical

structure to the logical constants, and remains agnostic about the meta-logical vocabulary: terms like 'is true' and 'is a consequence of'. According to him, this allows him to remain neutral on one of the questions mentioned above, namely the question of whether there is such a thing as 'the' consequence relation. It may be debated whether that is the case, but I will not take issue with it. It's the logical constants which are Sider's target, and they will also be my target.

IV. Sider's Arguments for Logical Realism

Why would one believe in logico-metaphysical structure? Sider has several arguments, which I've named the Argument from Semantic Determinacy, the Argument from Ontological Realism, and the Argument from Ideological Commitment.

The first argument takes us from metasemantic considerations to metaphysical conclusions. The argument from semantic determinacy considers threats of radical semantic indeterminacy with regard to the logical constants, particularly along the lines of the so-called 'Kripkenstein' problem. Logico-metaphysical structure is proposed as the factor that could explain why the logical constants are nevertheless determinate in meaning. As noted above, one of the theoretical roles that structure plays is that of supplying reference magnetism. If certain meanings for the logical constants are reference magnetic, then this can solve the indeterminacy problem. If certain meanings for the logical constants were structural, then they would be reference magnetic. And so, since the meanings of the logical constants clearly are determinate, they must be structural terms.

The argument from ontological realism brings meta-ontological considerations to bear on the matter. Sider argues that logico-metaphysical structure is a commitment of ontological realism, and the independent attractiveness of ontological realism then provides a motivation for believing in logico-metaphysical structure. Here's how the argument works, roughly – I'll go into much more detail in Chapter II.

Suppose the quantifiers were not structural. Then there would be no reference

magnets that helped determine their meanings. Then their meanings would be determined solely by how we use them. Then, if use of the quantifiers differed radically between different people, the meanings of the quantifiers would differ between these people as well. The use of the quantifiers does differ radically between people in some cases, namely in cases where philosophers disagree massively about what exists. That would mean, then, that the meanings of the quantifiers differ between these people. But then their disagreements about what exists would be verbal ones: they wouldn't really be talking about the same question. If ontological debates are verbal, then there isn't a real fact of the matter about what exists. But there clearly is a real fact of the matter about what exists. So the debate cannot be verbal, so the meanings of the quantifiers cannot differ between these philosophers. So, given that their use of the quantifiers differs massively, some reference magnetism must explain why their quantifiers mean the same thing. So the quantifiers must be structural terms. But the quantifiers are logical terms, so there must be logico-metaphysical structure.

These two arguments are going to be discussed at length in this thesis. I will talk about the argument from ontological realism in Chapter II. I argue in that chapter that the link between logical realism and ontological realism isn't as close as Sider makes it out to be. I will set out how an ontological realist can do without logico-metaphysical structure, and in fact I'll argue that they would be better off without it. The argument from semantic determinacy is discussed in Chapters I and IV. In Chapter I, I spell out the radical indeterminacy worry that Sider alludes to (he does not spell it out himself) and I show how logico-metaphysical structure could provide an answer. I also argue, however, that it need not be the only answer. In Chapter IV I set out an alternative that I think does better.

The third argument, the one I called the argument from ideological commitment, doesn't get a chapter of its own in the thesis. Instead I will discuss it here.

V. The 'Quinean' Argument from Ideological Commitment

Sider writes:

“Quine’s (1948) criterion for ontological commitment is good as far as it goes: believe in those entities that your best theory says exists. But in trying to decide how much structure there is in the world, I can think of no better strategy than this extension of Quine’s criterion: believe in as much structure as your best theory of the world posits. The structure posited by a theory corresponds to its primitive notions—its “ideology” in Quine’s (1951) terminology—which includes its logical notions as well as its predicates.”⁵

This is a very general and, if valid, powerful form of argument. It would license us to regard any vocabulary that proves indispensable or at least very useful in theorizing as capturing some part of the structure of reality. This form of argument may give us a good reason for regarding the meanings of logical connectives as part of reality. After all, they do seem to crop up in any theory of the world. But this sort of argument has limitations, which come out when we ask *why* we should adhere to a criterion like this.

What is implicit in this principle (and in what Sider calls ‘Quine’s criterion for ontological commitment’⁶) is an inference to the best explanation. We need to explain why it is that a certain theory has more success than its rivals: the best explanation is (the argument alleges) that it captures more of reality than its rivals. To make this inference, one needs to have a certain background commitment to realism. A thorough-going anti-realist will not accept that theoretical success needs an explanation in terms of what reality is like. This might be thought of as a limitation on the argument: it will only work for realists. I do not, however, have any

5. Sider (2009), p. 37. It should be noted that this isn’t actually quite the way that Quine uses the term “ideology”. He uses it to refer to all of the notions expressible in a theory, whether primitive or not. He refers to the primitive notions of the theory as the theory’s ‘absolute ideology’.

6. There is some scope for confusion here. The term ‘Quine’s criterion’ is quite often used to refer to a different, more specific, principle: that we find out the ontological commitments of a theory by looking at the quantifications that occur in a suitably regimented version of that theory.

bone to pick with that sort of general realist attitude, so I'll just concede the implicit realist premise.

However, the fact that the argument involves an inference to the best explanation also means that it is open to a general caveat: if, for some element e of a theory, we can argue that e bears no responsibility for the success of that theory, we have no warrant to conclude that e captures a bit of reality. And it may be argued that logical notions indeed bear no responsibility for the success of theories. The use of logical connectives is not at all a distinguishing mark of successful theories. Unsuccessful theories use them as well.⁷

Maybe that objection will be thought of as missing the point. The comparison (one might think) is not between theories that have logical connectives and theories that don't, but between theories that have certain useful meanings for their logical connectives and theories that have other less useful meanings for them. But by understanding the argument in that way, we do not make it come out any better. It seems that virtually all theories have the same meanings for their logical vocabulary, and that this does not differ along with the relative success of these theories. (Of course, few theories come along with an explicit semantics for their logical terms.) While there are many different views on the meanings of the logical constants, proponents of a particular semantics for the logical constants will generally apply that semantics across the board, wherever the logical constants crop up.⁸ Therefore there seems to be little scope for correlating different sets of meanings for logical terms with different rates of theoretical success.

This objection can be put another way, which I think gets us to the heart of the matter a little more. Even if we have a general reason for regarding the ideology of a successful theory as capturing the structure of reality, we still need a reason to think

7. Not all theories need to use logical constants. An example of a theory without logical structure would be Parmenidean monism. In its original formulation, in Greek, the theory is stated in a single word, and Parmenides indicates that this is the only possible formulation. I would not regard it as a clear example of a successful theory, though.

8. I simplify a little. One could imagine varieties of logical pluralism on which different theories are interpreted as using different logical constants. For example, one could think that most theories use constants that are to be interpreted classically, some theories (mathematical ones for instance) use constants that are to be interpreted intuitionistically. But on such a view the difference in logical constants would not be regarded as indicating a difference in rates of theoretical success, and it's hard to imagine a pluralist view on which that would be the case.

that logical terms are truly part of the ideology of a theory. In the quotation above, Sider simply assumes that they are. The fact that logical vocabulary is generally stable across theories isn't clearly evidence in favour of the view that these terms are part of theories' ideologies; in fact it could be counted as evidence *against* this view. It ought to be a live option that the logical vocabulary is a neutral apparatus for formulating theories, not properly forming part of those theories' contents. Perhaps Sider is right that we should believe in what our best theories of the world posit: but he hasn't made the case that the logical vocabulary is part of what our best theories (or any theories) posit.⁹

We can draw a more general moral about Sider's criterion of ideological commitment from these considerations. What gives the criterion its intuitive plausibility and attractiveness is an implicit identification of what ought to be two different notions. On the one hand we have the notion of a theoretical primitive, understood simply as a term that appears undefined in the formulation of a theory. On the other hand we have the notion of what we might call an ideological posit: a basic concept that plays a role in determining the content of the theory. It is plausible that the elements of a theory that help set it apart from other theories ought to be appealed to when we're trying to explain why that theory is as successful as it is. The ideological posits of a theory would therefore seem to be the right sorts of things to attach metaphysical significance to. But we can't get round to drawing metaphysical conclusions unless we have some way of identifying the ideological posits. It is at this point that one is tempted to look at the theoretical primitives of the theory. That makes sense: one is looking for basic notions, so one looks for undefined terms.

But at this point I think one ought to be a bit more careful than Sider is. It is fine, I think, to look for the ideological posits of a theory among its theoretical primitives. But it is not fine to just assume that every theoretical primitive will be an ideological posit. We simply have no prior guarantee that such an assumption will always be borne out. I think one ought to consider, with a given theoretical primitive, whether it plausibly plays a role in determining the content of that theory, and setting it

9. Similar problems arise for I.B.E. arguments for mathematical realism.

apart from other theories. This isn't something that one could simply read off from the formulation of the theory, in the way that one *can* read off whether a certain term is undefined or not. One may have to look at how these same terms appear in other theories, and one may have to bring some philosophy to bear on the matter. In the case of the logical constants, it is at the very least debatable whether these primitives play a role in determining the content of any theory (any theory in which they are used rather than mentioned, that is). I think that the view on which the logical constants are merely a device for formulating theories deserves a fair chance, and if that view is on the table then one would beg the question against it by going along with Sider's argument from ideological commitment. Sider's criterion of ideological commitment is too crude a tool, I believe.

So we should at least be wary of this argument of Sider's. But there is more we could do to counter it. A more decisive way of undermining the argument would be to give a precise explanation of why the logical vocabulary can be found in all of our theories, an explanation according to which this vocabulary does not play a role in helping us get at the structure of the world. If we had such an explanation, then there'd be no need to posit logico-metaphysical structure to account for this vocabulary, and Sider's Quinean argument would lose its force. That is the strategy I will take in Chapters III (for the logical connectives) and V (for the quantifiers).

VI. The Structure of the Thesis

In Chapter I, I set up and discuss the merits of Sider's argument from semantic determinacy. I start off by describing a general problem of radical semantic indeterminacy: the Kripkenstein problem. I then set out the basics of a Lewisian metasemantics (which Sider broadly subscribes to). I explain how, by introducing the notions of naturalness and eligibility, one can use this metasemantics to explain why the meanings of our words can be determinate after all. I then set out a version of the Kripkenstein problem for the logical connectives, and one for the quantifiers. I then describe how Sider's introduction of logico-metaphysical structure can stave off indeterminacy in the logical vocabulary in broadly the same way that naturalness does in the general case. I then raise some qualms about the notion of logico-

metaphysical structure. I conclude that we could block Sider's argument by coming up with a different explanation of the determinacy of the logical vocabulary which doesn't need to appeal to logico-metaphysical structure.

In Chapter II, I discuss Sider's argument from ontological realism. I start off by describing the long-running debate between Sider and Eli Hirsch on the status of ontological debates. I set out the structure of each of their arguments. I raise some qualms about Hirsch's argument, suggesting that Sider perhaps needn't go so far as he does to defend himself against Hirsch. I then raise a problem for Sider's externalist semantics for the quantifiers: on that view the validity of certain quantificational inference rules isn't a priori. I then discuss what would happen if Sider were to give up his appeal to logico-metaphysical structure for the quantifiers. I argue that ontological realism could still be defended, and that we don't immediately get the result that there is no fact of the matter about what really exists, even if ontological debates are verbal. Finally, I suggest out a semantics and metasemantics for ontological realists on which ontological debates don't come out as verbal, one that doesn't appeal to logico-metaphysical structure. In an appendix to chapter II, I examine what a proponent of this semantics could say about the question of unrestricted quantification.

In Chapter III, I set out logical expressivism for propositional logic. After first describing the ambitions and limitations of the project, I start by describing some important notions, in particular the propositional attitudes of acceptance and rejection, and the notion of ruling out a pattern of acceptances and rejections. I explain how the significance of uttering a logically complex sentence can be understood as that of expressing a complex propositional attitude: the ruling-out of a pattern of propositional attitudes.

I then set out a formal language, Sequentese, a language based on multiple-conclusion sequent calculus. It has no logical connectives but allows embedding of sequents within sequents. I use this language to generate expressions for complex propositional attitudes. I then give a method for translating a given sentence of propositional logic into a sequent of Sequentese (and vice versa), which can then in turn be translated into a description of a complex propositional attitude in English. I show a recoverability result for these translation manuals.

Finally, I describe a set of inference rules for Sequentese, which I motivate in terms of a small number of principles about synchronic theoretical rationality. I then show how the Hilbert axioms of classical propositional logic can be derived in the Sequentese calculus, and how the inference rules of Sequentese can in turn be proven in standard classical logic. I then wrap up with some philosophical considerations.

In Chapter IV, I return to the issue of semantic indeterminacy discussed in Chapter I, and bring the resources of logical expressivism to bear on the matter. I begin by setting out in more detail a Lewisian metasemantics, based primarily on his “Radical Interpretation” (1974). In this version of his metasemantics, the notion of an ideal interpreter plays a big role. I discuss how the notion of reference magnetism is to be understood within this metasemantic framework. I then adapt the Lewisian metasemantics to accommodate logical expressivism. One of the changes this involves is that the subjects of radical interpretation must now be understood, by the ideal interpreter, as creatures that have the sorts of complex propositional attitudes described in Chapter III. This in turn puts some constraints on the interpretation of their language: it must be a language suitable to the communication of complex propositional attitudes. In particular, this means that their language must include devices that allow for the constructions of sentences that express propositional attitudes of any given structure and complexity.

By picking out the terms that, on the basis of their use, can be accorded such a role, the radical interpreter isolates a set of 'merely expressive' terms, which can have their meanings explained in terms of the contribution they make to what sort of attitude the sentence expresses. These terms are the logical constants. Since, for these terms, the interpreter does not so much select the most likely candidate meanings that fit the use facts, but selects the most likely candidate terms that, on the basis of their use, fit the meaning, I call this mechanism 'reverse reference magnetism'. I then argue that this mechanism has the potential to stave off radical semantic indeterminacy for the logical connectives.

In Chapter V, I extend logical expressivism to the predicate logic. I discuss a few proposals for understanding the attitudes that might be involved in quantification, settling on a proposal derived from Ramsey (1929). On this proposal, universal

generalisations express general dispositions to attribute some property to any object one might encounter. To implement this idea in logical expressivism, I first introduce some new notions into the theory. I replace the attitudes of acceptance an atomic proposition and rejecting an atomic proposition with, respectively, accepting a predication and rejecting a predication. I also introduce the notion of a 'condition': the attitudinal counterpart of an open sentence. I use those notions to describe the attitudes involved in quantification.

I then extend the Sequentese formalism to cover these new attitudes, introducing 'open' and 'closed' sequents. I extend the translation manual and the inference rules, and I show that the resultant system is equivalent to classical predicate logic.

Finally, in the conclusion, I discuss the results of the foregoing chapters and consider what further challenges remain. I also discuss some possible applications of logical expressivism beyond the topics of this thesis. In particular, I discuss how one might use the resources of logical expressivism to formulate a version of moral expressivism that is not vulnerable to Frege-Geach problems.

Chapter I: The Argument from Semantic Determinacy

0. Introduction	15
I. The Indeterminacy of Meaning	17
I.I. The Kripkenstein Problem	18
II. Interpretation and Meaning	24
II.I. Lewis's Eligibility Response	25
III. A Kripkenstein Argument for Logical Constants	30
III.I. Kripkenstein Connectives	31
III.II. Kripkenstein Quantifiers	34
VI. Eligibility Responses	36
V. Conclusion	39

0. Introduction

In this chapter I will examine an argument for believing that the world comes equipped with what I'm calling 'logico-metaphysical structure'. This argument comes down, in the end, to an explanatory claim. If there is no logico-metaphysical structure, the argument says, we would be at a loss to explain something that needs explaining: the semantic determinacy of our logical terms. This argument, I'll concede, has some force. The datum that logico-metaphysical structure is supposed to explain is a genuine datum, and it is indeed not immediately obvious how it would be explained in the absence of logico-metaphysical structure.

But the force of that argument can be undermined. If the datum that needs explaining can be explained in some other way, without appealing to logico-

metaphysical structure, and if this other way of explaining is just as good (or better) in other respects, then we do not have a reason for believing in logico-metaphysical structure. This is what I'll argue is the case: there is another explanation available. A large chunk of this thesis is devoted to articulating a theory of logic which, among other things, provides that explanation. In this chapter, I will be setting up the explanandum. I will introduce some general issues and theories in metasemantics, explain the general puzzle of radical semantic indeterminacy, and articulate how that puzzle applies to the logical vocabulary.

The argument for logico-metaphysical structure which I'll discuss in this chapter is due to Ted Sider, as is the argument that chapter II is devoted to. Sider is one of few philosophers in recent years who have argued explicitly for a metaphysically robust view of logic. Sider believes that there has to be logico-metaphysical structure.¹⁰ The Siderian argument discussed in this chapter I call the argument from semantic determinacy. In the introduction we already discussed his argument from ideological commitment and in the next chapter we'll discuss his argument from ontological realism. As noted in the introduction, it may seem a bit lavish to devote so much space in a thesis to three arguments by the same author for a fairly unpopular view. But I believe it is justified, for the kinds of insights we can develop in arguing with Sider are of more general value. The goal is not to take pot-shots at Sider: the goal is to find out more about the metaphysics of logic, and the best way of doing that is engaging with the arguments of someone who has articulated views on the subject.

The chapter is structured as follows. In part I, I introduce the issue of radical semantic indeterminacy. Though there are several ways in which the puzzle has been articulated, I will present the one that Sider himself appeals to in motivating his views: the Kripkenstein puzzle. I discuss which theories are particularly vulnerable to this puzzle. I then explain how David Lewis's (1983, 1984, 1992) eligibility/naturalness response tries to tackle these problems, and discuss whether it does so successfully.

10. Note again that 'logico-metaphysical structure' is my own term. Sider just calls it 'logical structure'. I think that term has a little too much potential for creating confusion, so I want to avoid it.

In part II, I formulate a problem of semantic indeterminacy specifically for the logical connectives, in order to show that this part of our vocabulary is as open to these challenges as any other. I then explain how Sider's hypothesis of logico-metaphysical structure allows him to tackle this problem in the Lewisian manner explained earlier. I note some qualms about this metaphysical hypothesis.

At the end of the chapter I'll draw some tentative morals, and look ahead to the next chapter.

I. The Indeterminacy of Meaning

It seems undeniable that most of our words have fairly determinate meanings, and that there are right and wrong ways to use them. Moreover, we believe that we grasp these meanings, and know how to apply our words in lots of different situations. These are pretty established beliefs, and we would let it count against a theory of meaning if it claimed that we are, in fact, wrong about all of this.

Are we right to take this attitude? Why are we so certain that our words enjoy a fairly high degree of semantic determinacy? Might we be deceived about this? *Prima facie*, we would seem to have lots of evidence available for the determinacy thesis. After all, we evidently understand each other and we can explain what our words mean. But as I'll explain in a bit, there are sceptical challenges which seem to undercut all of this evidence. As with all sceptical challenges, there are then two attitudes to take. One is to give up the challenged thesis, at least for the time being, until more solid evidence comes in. In some cases, this may be the right thing to do. The other attitude is to stick to one's guns, and try to find something to challenge in the sceptic's argument. This latter attitude is the one that philosophers most often take nowadays to sceptical challenges, and it is the attitude that everyone has taken towards sceptical challenges concerning semantic determinacy.

Perhaps we generally take this attitude because we cannot imagine that anything the sceptic might appeal to in support of their argument could turn out to be more certain than our original belief in semantic determinacy. The belief in semantic

determinacy could be, by that token, a *Moorean belief*: a belief that is unshakeable because of the way our initial credences are distributed and the practices we have for revising our beliefs. Of course, presented in this way, it doesn't seem to be a particularly rational attitude to take: only if our initial confidence in the determinacy thesis was justified, would we be justified in sticking with it come what may. All I have to offer on this point, however, is the admission that I do in fact share the universal inclination to respond in this way to the sceptic: hold on to the belief in semantic determinacy, and look for a mistake in the sceptic's argument. And hope for the best.

As I just mentioned, compelling-seeming arguments suggest that in fact the meanings of our words are in fact not determinate at all. I will concentrate on one such argument, the famous 'Kripkenstein' problem articulated by Kripke in his (1982). Besides that problem, there is also Putnam's 'model-theoretic argument'. Closely related to these problems are Quine's 'Gavagai' problem of radical translation, and Goodman's 'new riddle of induction'.¹¹ I will concentrate on the Kripkenstein problem, for that is the problem that Sider appeals to in his argument from semantic determinacy.

I.I. The Kripkenstein Problem

The 'Kripkenstein' problem was popularised by Saul Kripke, who claimed to find it (or something like it) in Wittgenstein's *Philosophical Investigations*.¹² Since some have been reluctant to attribute the puzzle directly to Wittgenstein, a convention has sprung up of attributing it to a gerrymandered philosopher called "Kripkenstein". It is also known as the 'rule-following' problem.

The problem, to begin with, is that of assigning determinate meanings to someone's words on the basis of their use of them. Suppose we want to know if a

11. Putnam (1980), Quine (1960), Goodman (1979). It should be noted however that Goodman's riddle is in the first instance an epistemological puzzle, not a metasemantic one. It is not hard to figure out how to bring Goodman's reasoning to bear on metasemantics, however, and many have done so.

12. Kripke (ibid).

student is using the term “plus” to refer to the familiar mathematical operation of addition. (This example generalises widely: there is nothing significant about the mathematical nature of the example.) Well, the most obvious source of evidence to appeal to is their use of the term. We could give them a list of pairs of numbers, and ask them, for each pair, what the one *plus* the other would be. By answering these questions, they in effect rule out various interpretations of their words. If, for instance, by “plus” they meant what we usually mean by “minus”, then if we asked them “what is five plus three”, they would answer “two”. If they don't, then we know they don't mean by “plus” what we mean by “minus”. So we have a heuristic: observe their use of the term, and strike off candidate meanings that aren't consistent with that use, until we have narrowed it down to one candidate meaning.

Let's suppose they answer these questions, and it turns out that for each pair, what they give us is the sum of the two numbers. That's what we should expect, if they do use “plus” to mean addition. So it seems that the student is indeed using “plus” to mean addition.

But wait: the list of pairs of numbers will be finite, and so there will be some pairs of numbers that are not on it. We don't know at this point what the student would do with those numbers: perhaps they'll subtract one from the other, or multiply them. If that's what they would do, then they aren't using “plus” to mean addition at all: rather, they'd be using the term to pick out some function from numbers to numbers, one that looks like addition for the numbers that we've given them but not for some numbers that we haven't given them. So we have a problem: if our evidence of what their word means is observed use, and we're not in a position to observe all the possible uses, then we're not in a position to figure out precisely what the student is using the term to mean.

It seems we can't tell whether “plus”, as used by our subject, means **plus**, the function of addition, or whether it means **quus**, a function from pairs of numbers to numbers that differs from addition in some cases we haven't observed.¹³ Suppose for instance that the **quus** function (call it “quaddition”) takes us from the numbers 85 and 73 to the number 5, unlike the addition function, which would take us to

13. The term “quus” is Kripke's.

158. If we haven't asked the student what 85 plus 73 is, then we don't know that they wouldn't have answered "5". And so according to our heuristic set out above, we haven't narrowed down the student's meaning for "plus" enough yet.

So far, this looks like a problem about finding out what a term means: an epistemological issue. But really, the problem is a little deeper than that. For what *determines* what words mean, if not the way that they are used? There seems to be a metaphysical connection between use and meaning, besides the epistemic one. And so there also is a metaphysical problem about meaning here: if addition and quaddition are two different meanings and the student's use of "plus" is – so far – in accordance with both of them, then how is it determined metaphysically that "plus" means addition, and not quaddition? Of course, as the student goes on to use the word "plus" in more different situations, their use may rule out various quus-style meanings. But the problem is that there will always be more of those sorts of meanings, as long as the student's use of "plus" does not exhaust all of the possible situations where it can be used. Because there are so many possible situations where it can be used, there are going to be very many quus-style meanings that aren't ruled out. So the semantic underdetermination that seems to arise from this is quite radical.

The natural response to this problem is to look for other things, besides the use of the term, which might serve to narrow down its meaning. For instance, we might look not only at how the term is applied, but also at how it is described. We might ask the student to explain what they mean by "plus". We might reasonably think that the meaning of "plus" is not just determined by how it's used, but also by clarificatory claims that are made about the term. The student might just come out and say that they use the term to mean the function of addition, or that, in using "plus", they always try to supply the *sum* of the two numbers supplied.

Bringing that sort of information to bear on the determination of meanings would help a bit. We now know that the student's use of "plus" is connected, in a certain way, to their use of "addition" and "sum". But now we might raise a new, broader question. How do we know what set of meanings the student has for the complex of terms consisting in "plus", "addition" and "sum"? It could be that the student's word

“plus” means quaddition, that their word “addition” also means quaddition, and that the 'sum' of two numbers, as they use “sum”, is what you get when you quadd those numbers. This set of meanings isn't ruled out by the student's use of “plus”, “addition” and “sum”.

We can look further. Another plausible meaning-determining factor is how the student is *disposed* to use the term, in addition to how that actually have used it. Now, our only evidence of the student's disposition to use “plus” is their actual use of “plus”. So as far as our epistemic situation goes, an appeal to dispositions does not help us at all. But it may help with the metaphysical question about the meaning-determination of “plus”, because the student's dispositions to use “plus” may well be more rich than their actual use, in the sense that they have particular dispositions to use the term “plus” in a certain way in situations that they haven't yet been in.

Here things become a little bit more subtle. As noted before, as long as the student's use of “plus” did not exhaust all the situations where “plus” can be used, the use does not fully determine a meaning: there are always quus-style candidate meanings that differ from the 'standard' meaning of “plus” in exactly those situations that haven't been considered. Something similar would apply to dispositions, if the student's dispositions concerning “plus” do not exhaust all the possible situations where “plus” can be used. But maybe dispositions *can* be exhaustive in the way that use can never be. Maybe the student can have a disposition that does determine, for every possible situation in which “plus” can be used, a particular use of “plus”.

Call such a disposition an *exhaustive disposition*. Can there be such disposition? A dilemma now crops up regarding the way in which such a disposition might be 'implemented' in the student's mind. There seem to be two options. The first option is that the disposition is something like a list. The student has, in their mind, a separate representation of every single situation in which “plus” can be used, each twinned with a separate instruction for using “plus”. This looks implausible: it would seem to take cognitive resources that we finite being don't have. The second option is that the student has, represented in their mind, some kind of general recipe that,

when supplied with a possible situation where “plus” can be used, spits out a particular instruction for using the term. That seems more plausible, at a first pass.

On reflection, however, that is problematic too. For now the student doesn't have the instructions ready at hand: they need to *use* the recipe that's in their mind to get the particular instruction that they need for a particular situation. But consider what such an instruction might be: it might be, for instance, an instruction to take the two numbers and come out with the *sum* of them. Now a problem should become apparent. Rules or recipes are the sorts of things that need interpreting, before they can be used. Just as the student's belief that they meant the same thing by “plus” and “addition” did not ultimately help with the problem of determining the meaning of “plus”, so too linking up the student's word “plus” with a mental instruction-for-use doesn't really help. There will be quus-like interpretations of both “plus” and the instruction-for-use.

So the appeal to dispositions seems to run aground on this dilemma. Either the dispositions are in the form of an exhaustive list or they are in the form of a rule or instruction. It seems implausible that the student could have anything of the former kind; but something of the latter kind doesn't really help with the determination-problem.

There is another area in which we might look for information that can tell us about the meaning of the student's word “plus”, and for facts that can help determine what that meaning is: other people. Tyler Burge has argued persuasively that the facts that determine the meaning of a particular person's word cannot consist solely in facts about that person.¹⁴ In a nutshell, the reason is that we often regard people as misusing certain words, or as not quite grasping the meaning of them. What seems to be at work in such cases is that we regard the meanings of these words as determined by how they are used by a linguistic *community*, or perhaps more particularly by how they are used by certain experts in that community. A person can count as misusing a word if they do not use it in accordance with the way the community, or the experts in the community, use the term.

14. Burge (1979). See also Putnam (1975).

If we are persuaded by Burge's observations, and I think we ought to be, then that might provide an additional source of meaning-determining facts. We might look beyond just the student's use of "plus" and related words, and their dispositions to use those. We could look at all the uses the members of that student's community make of "plus". That's a lot more information, and it could serve to rule out all kinds of candidate meanings for "plus".

Our enthusiasm for this new source of meaning-determining facts should be short-lived, though. All the uses of "plus" of all the member of the linguistic community combined could still not amount to an exhaustive record of use for "plus", given that there are infinities of possible situations in which "plus" can be used, and all the communities' uses of "plus" can only add up to a finite set of usage facts. There will still be infinitely more quus-like meanings for "plus" that are consistent with all of those facts, so the radical semantic underdetermination we faced hasn't diminished in any significant sense.

There isn't anything special about the term "plus": the example generalises widely, and not just to function-words like "plus". With a little ingenuity, one can find ways to extend it to any name or predicate. So the problem of radical semantic underdetermination that it poses is quite general. What makes it even more paradoxical is that the metaphysical nature of the problem – that it applies not just to our knowledge of meaning, but to the determination of meaning – means that it applies in the first-person case as much as in any other. It may seem to me that my words have determinate meanings, but the sceptical problem suggests that my impressions are misleading.

Another aspect of the problem that makes it worrying is that it doesn't seem to appeal to any particularly controversial or theory-bound premises. The idea that the way words are used is an important factor in determining what they mean is a commonplace: it seems as undeniable as the original thesis that the meanings of our words are fairly determinate. We have seen one way in which one might work around that idea, though: we've tried, in a few ways, to exploit the idea that meaning of our words might be determined by more than just the way we use them. So far that strategy hasn't panned out, but there are some more radical

proposals we haven't tried yet.

The issues raised by Kripke in his (1982) have been the subject of much discussion, and what I've presented does not nearly exhaust those issues. There is a cluster of issues surrounding the so-called *normativity* of meaning. Kripke suggests that semantic facts have some normative consequences. If a word has certain meaning, then that doesn't just mean that it is used in a certain way, but that there is a way in which it ought to be used. It's been much discussed whether there is indeed such a normative dimension to meaning and what the force of the normativity would be if there is. It has been suggested that the normativity imposes additional difficult constraints on metasemantic theories.¹⁵

This sub-debate is not one which I will involve myself with, in this chapter or indeed in this thesis. It is the more straightforward underdetermination aspect of the Kripkenstein problem that Sider appeals to in his argument for logico-metaphysical structure. Logico-metaphysical structure is not intended to help with the possible issue of semantic normativity, and it doesn't seem to have any obvious application to it. So although it is an interesting issue in itself, it is less relevant to the task of assessing Sider's argument. I will therefore set it aside.

We now know how, in general, Kripkenstein-style scepticism about semantic determinacy gets off the ground. Ultimately I am not, in this thesis, concerned with the general form of the Kripkenstein problem. Rather, I'm interested in the problem insofar as it applies to the logical vocabulary, since it is that particular application of Kripkenstein scepticism that Sider appeals to in his argument from semantic determinacy. But before I can talk about Sider's views, I need to say more about general strategies for replying to Kripkenstein problems, because this will help us understand Sider's argument. I will set out the basics of a Lewisian metasemantics, and explain how, in the context of that theory, one can respond to the Kripkenstein sceptic.

II. Interpretation and Meaning

15. See, among many others, Boghossian (1989) and Hattiangadi (2006).

The Kripkenstein problem is a problem about meaning. It seems to make trouble for a very attractive metasemantic thesis: that the meanings of words are determined by their use (From this point on, I shall use the term “use” more broadly, to include *dispositions* to use). For it seems that meaning in fact outruns use: there are many meanings that our use is consistent with, and which therefore aren't ruled out as live candidates for being the meanings of our words. It looks, so far, as if the thesis that use determines meaning commits us to thinking that meaning is radically underdetermined.

Nevertheless, the view that meaning is determined by use is still attractive on other grounds: importantly, it offers the prospect of a reductive view of meaning (which is nice, because meaning facts just don't look like rock-bottom facts) and it also unifies the metaphysics of meaning with the epistemology of meaning (at least if one also, plausibly, holds that our epistemic access to meaning is through use). Last but not least, it is non-revisionary, in that it is guaranteed never to turn up results about what our words mean that are surprisingly at variance with how we use those words¹⁶. So it is worth trying to save this kind of view of meaning.

II.I. Lewis's Eligibility Response

David Lewis holds a metasemantic theory broadly of the kind just described – one on which use plays an important part in determining meaning. This is not surprising, for all metasemantic theorists do, in some way or other. More specifically, Lewis holds a kind of metasemantic view which, following Williams¹⁷, we may call ‘interpretationism’. Interpretationism is a view about the metaphysics of meaning which says that what it is for the terms of some language to have certain meanings just is for the *best semantic theory* of that language to assign those meanings to those terms. That means that if we want to maintain that meaning facts are

16. See Williams (2007), pp. 364-365.

17. Ibid.

objective facts, we must also maintain that there are objective facts about which semantic theory is the best one.

Consequently, much of the interpretationist's theoretical work goes into describing what features make a semantic theory a good one. One feature of good theories is that they fit well with the data. In the case of semantic theories, the facts about how words are used play the role of the data. It is in this way that the principle that use determines meaning is incorporated in Lewis's metasemantics.

Within this interpretationist framework, Lewis developed a unified response to the problems we have been discussing.¹⁸ In a very broad perspective, the response that Lewis gave was of the same variety as the ones we already tried above: he appealed to facts beyond the mere use of words. But in the details, it's a very different approach.

Lewis added to his interpretationist metasemantics the notion of *eligibility*, or suitability to serve as the meaning of a word. Eligibility is a characteristic that candidate meanings can have more or less of. Lewis does not treat all candidate meanings as if they are, to start with, equally likely candidates for being the meaning of the word we're interpreting. The deck is stacked in favour of some meanings, and against others.

Here is how that works, on Williams' interpretation of Lewis.¹⁹ As mentioned, which candidate meaning is the actual meaning of a word is determined by what the best semantic theory is for the language that the word is part of. The relative eligibility of meanings plays a role in determining which theory of meaning is the best one. As we've seen, one good-making feature of a semantic theory is that it assigns meanings to words that fit well with the use of those words. According to Lewis, another good-making feature is that the theory assigns relatively eligible meanings to words. The actual meaning of a term will then be the one predicted by that theory that strikes the best balance (the appropriate notion of balance is a matter of further theoretical fine-tuning) between the fit-with-use of that theory and the relative eligibility of the meanings that it ascribes to terms.

18. See his 1983, 1984, and 1992.

19. Williams (unpublished).

The most clear-cut cases where eligibility plays a useful role are the cases where it serves as a tie-breaker of sorts. Suppose we're in one of the many cases which we've encountered above, where the use facts about a word fail to narrow down the candidate meanings to one. If one of the candidate meanings is more eligible than the others, then a semantic theory that attributes that meaning to the word rather than one of the other candidates will be better for that reason. And so that meaning will then end up being the meaning of the term, since the best theory says so. So if we can expect addition to be a more eligible candidate meaning than quaddition, we'll get a determinate result after all, and moreover we'll get the right determinate result.

But while the most obvious theoretical role of eligibility is to serve as a tie-breaker in cases like these, where considerations of use don't serve to fix the meaning of a word, it's not ruled out by Lewis that in certain cases a candidate meaning for a word might be so eligible that it will 'trump' meanings that fit the use of that word better. It is worth noting that this latter possibility might potentially undermine the non-revisionist ambitions of the interpretationist view of meaning noted above²⁰. This possibility of trumping is also one that other philosophers, including Sider and Weatherson, have made grateful use of^{21, 22}.

Meanings are not brutally eligible. Lewis believes that meanings are eligible in virtue of having more or less of a certain objective and intrinsic feature. Lewis ties down his notion of semantic eligibility to a metaphysical distinction between *natural* and *unnatural* properties, a distinction that also performs a variety of other

20. Ibid., also Hawthorne (forthcoming).

21. See Sider 2001 and Weatherson (2003)

22. Note that the fact that some candidate meaning is very eligible does not mean that it is always in the frame to be the meaning of a given term. Given the use facts concerning a certain term, there will be a certain range of possible meanings that count as genuine candidates meanings of that word. These are meanings that fit the usage facts fairly well but not necessarily perfectly. It is only among those meanings that we look for the most eligible ones, when building semantic theories.

If we counted every meaning as a genuine candidate interpretation for every word, then the possibility of eligibility trumping imperfect fit with use could make for some strange results. Suppose, say, that Zsa Zsa Gabor were a really really eligible candidate meaning, and that we regarded her as a potential meaning in every case where we interpreted a word. Then for any word that we would try to assign a meaning to, we'd have to weigh the fit-with-usage against the eligibility of the candidate meaning **Zsa Zsa Gabor**. But if **Zsa Zsa** were eligible *enough*, then even if she fitted badly with the use of that word, she would come out as the meaning of that word. As a result, **Zsa Zsa Gabor** could turn out being the meaning of every word. Which would be strange.

jobs in his philosophy²³. Although Lewis thinks we might be best off treating this distinction as primitive, he explains it informally as follows: natural properties are those whose sharing makes for real (that is, objective) resemblance between objects. The natural properties are those that sort the things in the world into classes that display a genuine unity. Unnatural properties, on the other hand, make for artificial-seeming, gerrymandered classes of objects. Some things 'go together' naturally; others not so much.

It's the properties that are more natural that are more eligible, i.e. if that property were a candidate meaning for a certain word, it would be one of the better candidates. However, if naturalness is to serve as a determiner of eligibility, it ought to be the sort of thing that comes in degrees, because eligibility comes in degrees. Lewis accommodates this requirement by postulating a 'ground floor' of *perfectly* natural properties and proposing that we define, somehow, a relative notion of naturalness in terms of those properties.

Here is how that might go: imagine a language which has simple terms for all the perfectly natural properties, and the means for defining other property-terms in terms of those: a set of logical connectives, perhaps set-theoretic and mereological vocabulary. Now provide definitions for other, non-perfectly-natural properties. The more relatively natural properties are, the simpler their definitions would be in that language.

To illustrate, take Goodman's famous example of the colour green and the colour grue. We can define the colour green, let's suppose, in physical terms, in terms of a particular continuous range of wave-lengths of light. To define the colour grue, we need to do more than that: we need to first define both green and blue, and then define grue disjunctively in terms of them. The definition of grue would turn out to be more complex than the definition of green.

Whether this particular bit of Lewis's proposal works out is a matter of controversy, but for the moment I'll take it that some measure of naturalness can indeed be found, so that there is a meaningful notion of relative naturalness

23. Lewis (1983).

available.²⁴

When we think of eligibility as relative naturalness and we think of relative naturalness as measured by complexity of definition, we can tell a bit more of a story about why eligibility plays a role in deciding which semantic theories are better than others. There is a more general theoretical virtue of simplicity: theories are better, *ceteris paribus*, if they're simpler. A semantic theory that attributes relatively eligible and hence natural meanings to words is by that token a theory that employs less complex notions, because more eligible meanings have shorter definitions in terms of the perfectly natural properties. That semantic theory will therefore be simpler. And therefore it will be better, given that simplicity is a good-making feature. So this good-making feature of semantic theories, that of attributing more eligible meanings, is really an instance of the more general good-making feature of simplicity.²⁵

The eligibility of candidate meanings helps to determine the meanings of words in a way that's quite different from how the use facts help to determine the meanings of words. From the perspective of interpretationism, the underdetermination problem that the Kripkenstein sceptic points out is the problem that, however rich the data are that a semantic theory is supposed to have a good fit with, there will be many semantic theories that succeed in fitting with that data. As far as that theoretical virtue goes, many theories will be equally good: theories that interpret "plus" as meaning addition, and theories that interpret it as meaning quaddition. By saying that our interpretations of "plus" should also fit with dispositions to use "plus" or with other people's use of "plus" we don't address this problem. We enrich the set of data that semantic theories are supposed to fit with,

24. See Weatherson (draft) for some qualms. It's controversial what Lewis's exact view were on naturalness and eligibility: some argue (Weatherson (ibid.), Schwartz (draft)) that naturalness only factors into the Lewisian story of meaning in a very indirect way, by playing a role in his theory of mental content, which in turn plays a role in his theory of linguistic content. The view I've set out here, where naturalness also bears on meaning indirectly, by way of bearing on the relative goodness of semantic theories, is due to Williams (draft), and it is also contentious. Furthermore, doubts have been raised about whether Lewis can get away with having just one all-purpose notion of naturalness. Weatherson (ibid.) argues that the metasemantic applications of naturalness require a distinct notion, one that's not tied down to the microphysical in the way that Lewis's is.

25. Williams (unpublished).

but however much we enrich it, there will be many theories that are equally matched in this regard.

The appeal to eligibility avoids this pitfall. If only certain meanings are particularly natural and hence eligible, semantic theories that are evenly matched when it comes to fitting with the data are not evenly matched overall. If the function of addition is particularly natural, a semantic theory which involves that property will be simpler and better for that reason. On interpretationist principles, it will come out on top for that reason, and our words will have the meanings attributed to them by that theory. Of course, it needs to be the case that addition and not quaddition is particularly natural. We don't have a guarantee that this is the case, so Lewis's response to the Kripkenstein puzzle does involve a certain amount of optimism.

We have, in the Lewisian 'externalist' view of meaning determination, a promising response to the Kripkenstein worry. Note, however, that it requires us to make some metaphysical assumptions: we need to believe that some properties are objectively different from others, in being more natural than the others. We need to believe, in a Platonic phrase that has recently gained popularity again, that nature has 'joints'. Or, as Sider would put it, structure.²⁶ This is a metaphysical commitment. It is not one to which I particularly object. But the way in which Lewis's response to the Kripkenstein problem commits us to metaphysical structure is exactly the way that, according to Sider, the Kripkenstein problem for the logical constants commits us to logico-metaphysical structure. And that is something to which I do object.

III. Kripkenstein Arguments for Logical Constants

Now that we've introduced the general problem of Kripkenstein indeterminacy and explained how the Lewisian externalist response works, we can turn to Sider's argument from semantic determinacy. The argument is simple. Since the logical constants are just as open to worries about semantic indeterminacy as other terms, our metasemantics will have to include a story that explains how these terms end

26. Sider (2012).

up having determinate meanings. We know how such a story might go in general: we have Lewis's response to take as a model. Lewis's response requires us to appeal to some objective feature of the world that can render certain candidate meanings more eligible than others. If we want to use the Lewisian eligibility response in the case of the logical constants, then there has to be some feature in virtue of which certain candidate meanings for the logical constants are more eligible than others. So in Sider's phrase, we must posit 'logical joints in reality', i.e. logico-metaphysical structure.²⁷

Sider does not, however, describe exactly how problems of semantic indeterminacy arise for the logical constants. In this section, I'll set out an indeterminacy problem of the Kripkenstein variety, borrowing a page from Goodman's new riddle of induction.²⁸ I'll first provide one for the case of the logical connectives. In the section after that, I deal with the case of quantifiers, which bring some additional complications along with them.

III.I. Kripkenstein Connectives

One might think, at a first pass, that there could be no Kripkenstein problem for logical connectives. After all – the thinking might go – there are only a small number of meanings that a logical connective could have. Logical connectives are truth-functional: they represent functions that take, as arguments, the truth values of the propositions flanking or following the connective (depending on whether it is, respectively, dyadic or unary) and give out truth values as values. For dyadic connectives there are sixteen different such functions, and for unary connectives there are four. To establish the meaning of a logical connective, we would only have run through sixteen or four possible interpretations and opt for the one that is most apt. No significant indeterminacy could arise.

But the assumption of truth-functionality is too quick. We cannot establish that

27. Sider (2009), p. 25.

28. Goodman (1979).

on the basis of use. Consider a mystery proposition M , which by hypothesis has never been stated. Let us call any proposition M -involving if it either just is proposition M , or is a complex proposition with M as one of its atoms. Consider a function that takes two propositions as arguments, and gives out the value True iff both those propositions are true, but only when neither proposition is M -involving. When at least one proposition is M -involving, it gives out the value True iff at least one propositions flanking it is not true. Here's how that would look more formally (ϕ and ψ are variables ranging over propositions):

For all ϕ , for all ψ , ' $\phi \& \psi$ ' is T iff either

(ϕ is not M -involving and ψ is not M -involving) and (ϕ is true and ψ is true), or

(ϕ is M -involving or ψ is M -involving) and it is not the case that (ϕ is true and ψ is true).

How are we to decide whether our connective “&” represents the familiar function that gives out the value True iff both propositions flanking it are true, no matter which propositions they are (call this function “**and**”), or this more eccentric function?²⁹

Someone might object that proposition M must be a marvellous proposition, to make an ordinary truth-function behave in such an eccentric way. What proposition could exert such a mysterious force? Why are we supposed to accept its possibility? But this response gets things wrong. There need be nothing weird about proposition M - it could be perfectly ordinary. But there might be something weird about the meaning of our logical connective “&”. It may just have, for a meaning, a function that behaves eccentrically around the proposition M , through no fault of that proposition. The question is not: could there be a proposition M that makes truth-functions step out of line, but: how do we know that a connective expresses a truth-

29. I put “**and**” in bold to mark it out as a semantic value and distinguish it from ordinary uses of “and” in the course of my exposition.

function to begin with?

The eccentric meaning for “&”, call it **quand**, seems gerrymandered. One might think that there is a presumption in favour of plain old truth-functional **and**, given its comparative simplicity. But this is an accident of what we take as primitive. In the same way that “green” can be defined in terms of “grue” and “bleen”, we could have a primitive connective “quand” that expresses **quand**, and give “&” its familiar meaning (**and**) in terms of that. Here's how that might go (“ ϕ ” and “ ψ ” are variables ranging over propositions):

For all ϕ , for all ψ , ‘ ϕ & ψ ’ is true iff either

(ϕ is not M-involving and ψ is not M-involving) and (ϕ is true quand ψ is true), or

(ϕ is M-involving and/or ψ is M-involving) and it is not the case that (ϕ is true quand ψ is true).

This makes **and** just as gerrymandered a meaning for “&” as **quand** seemed before. But what might seem suspicious is that the definition employs the English “and” on the right-hand side. Does this “and” express **and** or **quand**? If the former, then we will only have succeeded in showing that we can define **and** in a gerrymandered-looking way by using both “quand” and “and”. That's not quite what we were aiming for. But if we replace the “and”s with “quand”s, does the definition still come out correctly? It would depend on whether there were any M-involving propositions in the definition.

The difficulty can be averted, with some ingenuity. Informally speaking, the function **quand** behaves like Boolean **NAND** (i.e. $\neg[P \text{ and } Q]$) when it has at least one M-involving proposition on its flanks. So if we make sure that all the occurrences of “quand” in the definiens have M-involving flanking propositions, we can formulate the definition as if we were using a truth-functional **NAND**-expressing connective. A first step is replacing, in the definiens, occurrences of “ ϕ ” and “ ψ ” with, respectively, the following logically equivalent but M-involving phrases:

“(((M quand M) quand M) quand ϕ) quand (((M quand M) quand M) quand ϕ)”

“(((M quand M) quand M) quand ψ) quand (((M quand M) quand M) quand ψ)”

These are conjunctions of respectively ϕ and ψ with tautological formulae built up from proposition M, where “quand” in its role of **NAND**-expressing connective does all the connecting work. Then, we use these ϕ - and ψ -equivalent phrases (abbreviating them as “ ϕ^* ” and “ ψ^* ”) to give the following definition of “&”, using “quand”-as-**NAND** as a connective:

For all ϕ , for all ψ , ‘ ϕ & ψ ’ is true iff:

((ϕ^* is true quand ψ^* is true) quand (ϕ^* is true quand ψ^* is true)).

Here we have a definition of “&”-as-**and** which uses only “quand” for a connective. For the sake of saving space I have not replaced the “iff” by an equivalent “quand”-construction, but this could be done as well. Note that “M” can be regarded as a schematic letter for any proposition that we've never encountered, so that we could generate any number of **quand**-like meanings for “&”. And also, following roughly the same recipe, we can generate qu-deviants for other logical connectives. This shows that the gerrymanderedness of **quand** and its kin is only an artefact of what we take as primitive.

We have the makings, then, of an indeterminacy problem for logical connectives. How about quantifiers?

III.II. Kripkenstein Quantifiers

Quantifiers are logical constants just as much as logical connectives are, or that at least is the common view. It may not be obvious at this point that the Kripkenstein argument extends to quantifiers as well. But in fact it does, and I'll explain how.

Let **quall** be a gruesome meaning for the universal quantifier " \forall ", along the following lines. It behaves normally in most cases, but when it occurs in a sentence which involves a mystery predicate X (even if X is only mentioned, not used) it behaves like an existential quantifier. " \forall " can be defined as meaning **quall** as follows, using " $\Phi(x)$ " as a variable ranging over open sentences:

For all x and for all $\Phi(x)$, ' $\forall x\Phi(x)$ ' is true iff either:
 Φ does not involve X and for any β , ' $\Phi(\beta/x)$ ' is true, or
 Φ involves X and for some β , ' $\Phi(\beta/x)$ ' is true.

Quall behaves like our familiar existential quantifier just when it binds an X-involving open sentence, and like our familiar universal quantifier otherwise. Let " $\#$ " be a quantifier that expresses **quall**. " \forall " can then be defined in its familiar meaning as follows:

For all x and for all $\Phi(x)$, ' $\forall x\Phi(x)$ ' is true iff either:
 Φ does not involve X and ' $\#x\Phi(x)$ ' is true, or
 Φ involves X and ' $\neg\# \neg\Phi(x)$ ' is true.

These definitions are symmetric in structure. So while **quall** looks pretty weird, this weirdness is not a matter of non-language-relative complexity, for the regular universal quantifier-meaning (call it "**all**") seems just as weird if defined in a " $\#$ "-based language.

Actually, we haven't quite shown that yet. The above definitions of " \forall " are both of the overall form of universal quantifications over sentences. I've chosen to do it that way, rather than in schematic form, to make explicit the quantificational

character of definitions. If they were in schematic form, they'd still be implicitly quantificational, but one might not pick up on it quite as clearly. The definitions above use English quantificational constructions. Do these correspond to our familiar universal quantifier? If they do, then the latter definition of “ \forall ” uses the familiar universal quantifier to define that same quantifier. We won't have properly shown, then, that we can define “ \forall ” as meaning **all** by using just the **quall**-expressing “ $\#$ ”. We would just have shown that we can define it in terms of both **all** and **quall**, which is far less interesting.

Here is a way to remedy this: in addition to our object-language level quantifier “ $\#$ ”, we introduce a term “quall” into our meta-language which represents a meta-language-analogue of **quall**. Here's how it could then be made to work:

“ $X(a) \rightarrow X(a)$ ” is true and it is not the case that for quall x and for quall $\Phi(x)$, it is not the case that ($\forall x\Phi(x)$) is true iff ($\neg\#x\neg\Phi(x)$) is true).

The insertion of an X -involving tautology makes sure that “quall” and “ $\#$ ” will be behaving as our familiar existential quantifiers. We then exploit the duality of the existential and universal quantifier to get a definition which is, in effect, a universal quantification.

If we have not encountered X , as we're supposing, we could not decide on the basis of past use whether the meaning of our symbol “ \forall ” is all or quall. Since both symbols appear equally gruesome if defined in terms of the other, straightforward considerations of simplicity don't decide the matter either. If there are no further facts to appeal to, there is a Kripkenstein problem here as much as there was with “ $\&$ ”.

It seems, then, that there's at least as much a cause to appeal to eligibility in the case of the quantifiers as there is in the case of the connectives, if we want to stave off Kripkenstein indeterminacy.

IV. Eligibility Responses

Introducing eligibility looks like a promising way of securing determinacy for logical connectives. As I've argued, neither considerations of fit with use nor considerations of simplicity will rule out the Kripkenstein variants on **and** and **all**. Some additional constraint has to be imposed on the interpretation of “&” and “ \forall ” We might propose, then, that certain meanings for the logical connectives are more eligible than others. Certain ones stand out: it is objectively the case that a semantic theory for the logical vocabulary would be better for including those meanings than others, *ceteris paribus*. Because **and** and **all** are eligible, and **quand** and **quall** aren't, we are better off interpreting '&' in terms of the former.

However, when we explained the basic Lewisian response to the indeterminacy worries in section II, we saw that Lewis tied down the notion of eligibility to a metaphysical distinction between perfectly natural and less-than-perfectly natural properties. Since the meanings of the logical connectives are not properties, we can't simply apply Lewis's solution as-is. Ted Sider therefore suggests extending the notion of naturalness beyond just properties, and let it apply to the sorts of things that are the meanings of logical connectives. In other words, he hypothesises that the world has an additional sort of structure, above the sort that Lewis endorses. Besides the normal 'joints' in nature that correspond to natural properties, there are logical joints as well.³⁰ That would be logico-metaphysical structure.

Sider doesn't do a great deal to clarify what he takes this sort of structure to be, and perhaps there's not much *can* say about it. It is simply that feature of reality which makes the case that there are more and less eligible meanings for logical terms. In his (2012) he makes explicit that he doesn't, in general, regard the notion of structure as one that can be explained in any way other than by indicating the theoretical work that it does. However, in the case of logico-metaphysical structure, all we know about the theoretical work it does is that it helps to determine the meaning of the logical constants, and helps explain why the logical constants are in

30. Sider (2009a) and (2012), in particular chapter 1.

our vocabulary.³¹ Nevertheless, it does seem clear that Sider does not intend to identify logico-metaphysical structure (logical structure, as he calls it) with some structure that we already countenance. It seems to be an additional form of fundamental structure.

But unspecific as this metaphysical commitment may be, it gives rise to some philosophical qualms.

First, let's say there's some fundamental feature of reality which helps to fix the meaning of '&'. Let's say there's also a feature of reality which helps to fix the meaning of '¬'. Then the conjunction and negation signs will have determinate meanings. Given that these two have determinate meanings, all of the sentential connectives can be given determinate meanings (assuming, for the sake of argument, that we'll be interpreting them classically), because these two are a so-called 'minimally complete' set of operators. So to fix the meanings of all the sentential connectives, we need only assume that there are meaning-fixing features of reality for '&' and '¬'.

However, things would have worked out the same if we had assumed instead that there had been meaning-fixing structure corresponding to '→' and '¬', or just for the Sheffer stroke. Those are also minimally complete sets. So which metaphysical assumptions should we make? Or should we regard it as indeterminate which minimally complete set has dedicated meaning-fixers out there in reality? Should we assume that all of the connectives correspond to bits of the structure of reality, even though some of those bits would be redundant? Is there some way of describing what logico-metaphysical structure is that avoids these awkward questions? Sider is aware of this worry, and concedes that the fact that his view raises this awkward question does not speak in its favour. But he bites the bullet, picking the 'redundant structure' option.³²

Second, logico-metaphysical structure, whatever it may be exactly, is an addition

31. In the next chapter we will see a further role that some of the logical vocabulary (the quantificational bit) can play. But it will not do much to illuminate the general notion of logico-metaphysical structure.

32. Sider (2012), § 10.2. The logical realist who is not a compositional nihilist might look for a companion in guilt in mereology: whether to treat "part" or "proper part" as fundamental seems a likewise arbitrary choice.

to and complication of our overall metaphysics. Positing such structure is a cost. It may be a cost that's worth paying, of course, if the only alternative is to allow that the meanings of our logical constants are radically underdetermined. I will argue that it is not the only alternative. In Chapter III I outline an 'expressivist' view of the semantics of the logical connectives, which does not involve an appeal to logico-metaphysical structure. But as I argue in Chapter IV, that view of logic can still be used to give a broadly Lewisian response to problems of semantic indeterminacy.

V. Conclusion

In this chapter I've discussed Sider's argument from semantic determinacy. I haven't yet argued against it. To do that, I'll need to set out my own positive view of the metaphysics of logic. That's what I'll do in chapters 3 through 5. In those chapters, I'll propose a theory of the logical vocabulary that serves to meet two explanatory requirements. First, in response to Sider's argument from ideological commitment (see the Introduction) it explains why the logical constants are part of our vocabulary. Second, in response to the argument from semantic determinacy, it explains how the meanings of the logical constants are rendered reasonably determinate. Chapters III and V will be mostly concerned with the former task, and chapter IV with the latter, though these explanations will turn out to be closely linked.

The Lewisian approach to metasemantics, some outlines of which I've already given in this chapter, will be a constant theme in this thesis. A major reason for this is that this is also Ted Sider's approach to metasemantics, and it would be best to engage with Sider's arguments on his home ground, so to say. But another reason is that I regard this metasemantic approach as promising and fruitful, and moreover, as amenable to being tinkered with in the light of my own theoretical commitments. If time and space were no object, it would be best if I could either provide the reader with independent arguments for adopting the Lewisian approach, or present my own proposals in terms of other metasemantic approaches as well. Time and

space is an object, however, so what I'll do is simply adopt the Lewisian metasemantics, without having argued for it in any detail. The upshot may be that the conclusions of the thesis are to some extent conditional on the correctness of that metasemantic approach.

Before I start on the task of constructing a positive view to set against Sider's, I'll first complete the task of setting his arguments for believing in logico-metaphysical structure. In the next chapter, I present his argument from ontological realism, which argues that only by believing in at least some logico-metaphysical structure (enough to secure determinate meanings for the quantifiers) can we maintain common-sense ontological realism: the view that questions about what there is have objective and language-independent answers.

Chapter II: Quantifiers and Metaontology

0. Introduction	41
I. Quantifiers and the Realism-Deflationism Debate	43
II. The Meaning of “\exists”	44
III. The Structure of the Arguments	47
IV. Hirsch's Argument: a Closer Look	50
V. Eligibility and Apriority	54
VI. The Sider-Hirsch Dialectic	60
VI.I. What Ought we Disagree About?	61
VI.II. Verbal Debates about Real Facts	63
VII. Semantics and Metasemantics	64
VII.I. Local and Global Quantifier Meanings	65
VII.II. Domain Assignments	69
VIII. Conclusion	70
IX. Appendix: Unrestricted Quantification	71

0. Introduction

In the last chapter we discussed an argument for believing in logico-metaphysical structure, put forward by Ted Sider. This chapter discusses a different argument, also due to Sider, which I'll call the Argument from Ontological Realism. According to that argument, we can maintain a (presumed to be common-sense) ontological realism only by believing in at least some logico-metaphysical structure, to be precise some logico-metaphysical structure corresponding to the quantificational

fragment of our logical vocabulary. I will be arguing that ontological realism, while it may come with some metaphysical commitments, doesn't in fact come with a commitment to logico-metaphysical structure. If there is a real commitment, it is to a privileged domain of quantification, and not to a privileged quantifier meaning.

Of course, if ontological realism *did* commit us to believing in logico-metaphysical structure, it would only strictly commit us to believing in that bit of logico-metaphysical structure that corresponds to the quantifiers. It would take some additional argument to extend that commitment to the rest of the ontological vocabulary. This means that there would be a natural fall-back position available: concede the case of the quantifiers, but not the case of the connectives. But I'm confident it won't come to that: the argument from ontological realism can be resisted.

The structure of this chapter is as follows:

In section I, I explain the outlines of the debate between ontological deflationists and ontological realists, and explain how it impinges on the question of logico-metaphysical structure.

In section II, I'll discuss and compare Hirsch's and Sider's views on the metasemantics of quantifiers. I'll explain how their metasemantic views shore up their respective views about ontology.

In section III, I set out what I take to be the structure of Hirsch's and Sider's arguments, and identify key moves and premises that I want to challenge. This will set up the various topics discussed in the remainder of the chapter.

In section IV, I discuss Hirsch's case for ontological deflationism. Hirsch claims that interpretative charity bids us interpret ontological debates as merely verbal. I argue that the case is not nearly as compelling as Hirsch presents it as being.

In section V, I raise a problem for Sider's articulation of ontological realism: it gives the unexpected result that the validity of existential generalisation is not *a priori*. To avoid the result, it seems we must give up a key claim that shores up ontological realism.

In section VI, I discuss whether ontological realism is salvageable. I analyse the

realism-deflationism dialectic, arguing that the realist need not in fact defend their meta-metaphysical position on the basis of prior metasemantic claims, contrary to the impression one would get from the Sider-Hirsch exchange. I explain how one might defend realism in a much more straightforward, even flat-footed way.

In section VII, I argue that by drawing a useful distinction at the level of semantics – a distinction between different notions of quantifier meanings – we see that a more attractive realist metasemantics for quantifiers is available, which lacks the worrisome aspects that Sider's has.

In VIII, I return to Hirsch's arguments for ontological deflationism, and point out some weak points in his case.

In section VIII, I wrap up and look ahead to the next chapter.

At the end of the chapter, an appendix is included which discusses an issue ancillary to the main argument: the challenge of securing the possibility of unrestricted quantification. I set out a view of this issue which is consonant with the form of ontological realism I put forward.

I. Quantifiers and the Realism-Deflationism Debate

For about a decade, Ted Sider and Eli Hirsch have been debating whether ontological disputes are substantive. Suppose I have a matchbox with two matches left in it. How many things are in it? Two (the two matches) or three (the two matches and their 'sum', which has the two matches as parts)? Hirsch's position, which takes off from Carnap and Putnam, is that in such cases there is no language-independent fact of the matter. Different languages describe the contents of the world differently, and no language gets it metaphysically more right than others. And since there's nothing to choose between differently quantifying languages, we should stick with the one we have, and settle ontological disputes by analysing ordinary language.³³ Against Hirsch, Sider defends 'ontological realism': there is one

33. See Carnap 1950, Putnam 1983, Hirsch 2002, 2004, 2005, 2008, 2009a.

way of counting the things in the matchbox that gets things metaphysically right, and metaphysicians are tasked with finding out which way that is.³⁴

What makes this debate between Sider and Hirsch relevant to the question of logico-metaphysical structure is the way that Sider defends his ontological realism. He defends his claim that there is an ontologically privileged language with an appeal to eligibility. There is one meaning for the quantifier that is particularly eligible when we are speaking philosophically about what there is, he suggests. When the existential quantifier of our language has that meaning, then by speaking truly in our language we would be saying what there really is. But why is that meaning so eligible? Because it's *natural*, in the sense explained in chapter one: it's part of the structure of reality. So in short, Sider makes a claim about logico-metaphysical structure – about quantificational structure, more specifically – which shores up his ontological realism. It seems then that what side we take in the debate between realism and deflationism may affect our views on logico-metaphysical structure.

I am on Sider's side on the general question of ontological realism. I think it's a more attractive and common-sense view than deflationism. What is more, Sider seems (admittedly on the basis of anecdotal evidence) to have most metaphysicians on his side. If Sider can make the case that a commitment to logico-metaphysical structure is something that simply comes along with ontological realism, then for many metaphysicians that would be a good reason to believe in quantificational structure. In this chapter, I won't primarily be concerned with arguing for or against realism (although I will offer some reasons for thinking that Hirsch's arguments lack force). I'll mainly be concerned to question the link between ontological realism and logico-metaphysical structure. The former can be defended without appealing to the latter, I'll argue.

II. The Meaning of “ \exists ”

34. See Sider 2001a, 2001b, 2005, 2009a, 2009b, 2011 (Ch. 9).

Much of the meta-metaphysical argument between Sider and Hirsch turns on what determines the meaning of “ \exists ”, the existential quantifier, and its natural language counterparts: words like “there is” and “exists”.³⁵ Hirsch argues that if we are charitable to what disagreeing ontologists say about what there is, we should interpret them as attaching different meanings to “ \exists ”. By that token, their disagreement would be merely verbal. Sider argues that, while considerations of charity might indeed make us think that “ \exists ” is used with different meanings by metaphysicians of different stripes, these are not the only relevant considerations. The world comes ‘ready-made’, so to say, with a supremely eligible candidate meaning for “ \exists ”. This is the one that would have us use “ \exists ” in accordance with what really exists, out there. In philosophical disputes about ontology – where what’s really out there is our sole concern – the eligibility of this quantifier meaning trumps considerations of charity. So metaphysicians who disagree about matchbox contents should be interpreted as using “ \exists ” with the same meaning, and as disagreeing substantively.³⁶

Sider and Hirsch disagree about what determines quantifier meanings. They both believe the meaning of “ \exists ” is to an extent fixed by its logical inferential role, i.e. what we capture by the introduction and elimination rules. They also believe that more factors into it. Hirsch proposes that various extra-logical claims about existence help to determine the meaning of a quantifier. For example, the quantifications of a mereological universalist – someone who believes that for any collection of concrete objects there is a further object that has those objects as parts – should be interpreted to conform to the following principle (Hirsch’s own example).

“[A]ny sentence of the form “There exists something composed of the *F*-thing and the *G*-thing” is true if the expression “the *F*-thing” refers to

35. For simplicity's sake, I will concentrate just on the symbol “ \exists ”. Everything I say should be taken to extend to its natural language counterparts as well.

36. I suppress some inessential complications. In earlier work (2001a) Sider claims that in the case of the English “exists”, eligibility trumps charity, so that it is true in English to say of all and only the things that really exist, that they exist. Later (2004) he admits doubts about whether eligibility trumps charity in the case of English, but engineers a fall-back position: the proper use of the English ‘exists’ might not accord with the truth in ontology, but this will hold for the ‘exists’ of *ontologese*, the language that ontologists in the ontology room are trying (implicitly) to speak.

something and the expression “the *G*-thing” refers to something.”³⁷

But a mereological restrictivist – who believes that concrete objects don’t always compose further objects – wouldn’t profess a principle like that, so their uses of “ \exists ” shouldn’t be interpreted to accord with it. We thereby get the result that mereological universalists and mereological restrictivists use “ \exists ” with a different meaning and disagree merely verbally about the matchbox’s contents.

(At this point it might be worth noting that Hirsch, in setting up his argument, makes a certain idealising assumption. Hirsch is worried, for reasons to do with Tyler Burge's (1979) arguments for social externalism, that it makes no sense to interpret a single ontologist's utterances in isolation, since any individual inherits their language from the community they are part of. For that reason, Hirsch has us imagine that these ontological debates are held between members of different 'tribes', which have different practices as regards quantification. So there is a tribe of mereological universalists, for which the principle quoted above is a platitude, and there is a tribe of mereological restrictivists, for which the principle above is trivially false. Although this idealising assumption does help us steer clear of the worries Burge's arguments would otherwise raise, we'll see in section IV that it's not entirely harmless.)

Sider agrees with Hirsch that the formal inferential role of “ \exists ” helps fix its meaning, and also that extra-logical claims about what exists play a meaning-fixing role. But he also thinks that there is another crucial meaning-determining factor: logico-metaphysical structure. He writes:

“The inferential role played by the symbol “ \exists ” in our use of that symbol is only part of what secures its meaning. Another part is the intrinsic eligibility of candidate meanings. Suppose that the world comes equipped with ‘logical joints’ as well as extra-logical ones. In particular, in addition to there being distinguished classes of objects that count as genuinely similar, the world comes ‘ready-made’ with a single domain **D** of objects: the class of all the objects there are. This class is the most

37. Hirsch 2002, p. 54. The context makes clear this is restricted to concreta.

eligible meaning possible for any symbol playing the inferential role of the unrestricted existential quantifier.”³⁸

Unless charity to ontological claims trumps the eligibility of **D**, the proper use of “ \exists ” reflects what there is in **D**.³⁹ In terms of our earlier example: the sum of the two matches either is in **D** or it isn't; consequently, either the universalist or the restrictivist is wrong about the matchbox's contents.

For Sider, *contra* Hirsch, it's the world (specifically, a bit of logico-metaphysical structure) and not merely our use of “ \exists ” which determines what “ \exists ” means. That Sider and Hirsch disagree about what fixes quantifier meanings doesn't mean, incidentally, that they are talking past each other. Rather, their debate about the substantiveness of ontology takes the form of a debate about the metasemantics of quantifiers, and they simply disagree about what determines quantifier meanings.⁴⁰

III. The Structure of the Arguments

To make clear exactly where one might intervene in the Sider-Hirsch debate, it may help to set out the structure of their arguments a bit more formally. Here's how I take Hirsch's argument to work:

(i) Philosophers with different views on ontology use “ \exists ” in very different ways. In particular, they endorse different general claims about what sorts of things exist. (Premise)

(ii) When we interpret the symbol “ \exists ” we do so by finding the most charitable interpretation: one that makes the assertions of language-users come out most rational/true. (Premise)

(iii) Interpreting philosophers with different views on ontology as using “ \exists ” with the same meaning would be uncharitable, given that one of

38. Sider (2001a), p. xxii.

39. Sider allows that this might happen, but not in the ontology room: see footnote 36.

40. But cf. Marsh (2009).

them must then be massively wrong. (from (i))

(iv) When philosophers with different views on ontology use the symbol “ \exists ”, the symbol has a different meaning as used by each of them. (From (ii) and (iii))

(v) A disagreement about the application of some term is verbal if the term has different meanings in the mouths of different parties in the disagreement, in such a way that their applications of it can be consistent with each other. (Partial definition)

(vi) Philosophers with different views on ontology disagree merely verbally. (From (iv) and (v))

(vii) Iff debates about a certain subject matter are merely verbal, there are no non-language-relative facts about that subject matter.⁴¹ (Premise)

(viii) There are no non-language-relative facts about ontology. (From (vi) and (vii))

The common ground between Sider and Hirsch, as I interpret their debate, consists in premises (i) and (vii) and definition (v). Sider has also on occasion conceded (iii). Sider disagrees with (ii), a claim about metasemantics, and by supplying some further premises of his own, arrives at a conclusion diametrically opposed to Hirsch's. Here is how I take Sider's argument to go:

(a) Philosophers with different views on ontology use “ \exists ” in very different ways. In particular, they endorse different general claims about what sorts of things exist. (Premise)

(b) When we interpret the symbol “ \exists ” we do so by finding the meaning that strikes the best balance between charity and eligibility. (Premise)

(c) Interpreting philosophers with different views on ontology as using “ \exists ” with the same meaning would be uncharitable, given that one of

41. I am taking it that disputes are either verbal or substantive, with no third option.

them must then be massively wrong. (from (a))

(d) There is a supremely eligible meaning-candidate for “ \exists ”, and it is the one that would make it true to say, of all and only the things that exist (as a matter of extra-linguistic fact), that they exist. (Premise)

(e) This meaning-candidate for “ \exists ” is so eligible that this would compensate for any deficits in charity. (Premise)

(f) Philosophers with different views on ontology use the symbol “ \exists ” with the same meaning. (From (b), (c), (d) and (e))

(g) A disagreement about a question is verbal iff the parties in the disagreement use terms in the question with different meanings, in such a way that their answers are not inconsistent. (Definition)

(h) Philosophers with different views on ontology do not disagree verbally. (From (f) and (g))

(i) Iff debates about a certain subject matter are merely verbal, there are no non-language-relative facts about that subject matter. (Premise)

(j) There are non-language-relative facts about ontology. (From (h) and (i))

This is one way to reconstruct Sider's argument. The lines in the argument might be rearranged, however, to make up different arguments, and depending on what articles of Sider's one looks at, one would get a different impression of what he's arguing from and towards. The argument presented here is, I think, the most natural interpretation of his (2009a). But for his (2001a), a more natural reconstruction would make (j) a premise, and (d) and (e) conclusions. And in his (2009b), (f) seems to be a premise, and (d), (e), (h) and (j) are conclusions. It does not matter a great deal, however, given that Sider makes largely the same set of individual claims, across his contributions to the Sider-Hirsch debate.

I'm going to disagree with a number of claims of both Hirsch and Sider. I'll disagree with Hirsch's (ii) and go along with Sider's (b), not only because I happen to

agree with Sider on this point but also because I'm most concerned in this chapter to argue against Sider, and it would be dialectically inappropriate not to go along with as many of his premises as I can. I will argue, in section IV, that (iii)/(c) doesn't follow if we carefully consider the nature of the ontological debate and plausibly fine-tune the notion of charity. I will argue in section V that Sider's premise (d) is too strong: with that premise in place, Sider's argument would have unwelcome consequences, in particular the result that the validity of some inference rules for "∃" turns out not to be an a priori matter. In section VI, I will argue that (vii)/(i) is simply false: a disagreement about some subject matter being verbal is consistent with the existence of non-language-relative facts about that subject matter. In section VII I will argue that if we draw some overlooked distinctions about the semantics of quantifiers, we can still get to conclusion (j) without going through (d).

IV. Hirsch's Argument: a Closer Look

As has been mentioned, Hirsch's arguments for ontological deflationism turn on the notion of interpretative charity. Charity, Hirsch says, obliges us to make people's utterances rational⁴²; it is in order to maximise rationality that we try to make people's utterance's come out true. But although Hirsch often seems to treat it as such, the number of true beliefs that we attribute to a person is not at all a straightforward measure of how rational the beliefs are that we attribute to them. Suppose that a subject S comes to believe, through some genuine failure of rationality, a very general a priori falsehood U, which is (let us suppose) that simples always compose. Let's suppose for the sake of argument that ontological nihilism is the truth of the matter. From U, together with ordinary empirical inputs, a vast number of falsehoods follow: that there is a table here, a person there, an Eiffel tower there, and a table-person-Eiffel tower sum scattered over all those locations. While being irrational in believing U, S will not become more irrational by believing these falsehoods; once we have attributed the mistake of believing U to S, charity to

42. Hirsch (2009), p. 238.

use, understood as an obligation to maximise rationality, obliges us to attribute all these false beliefs to S. By attributing to S true beliefs to the effect that there are no tables and chairs, we'd be making S less rational: those beliefs wouldn't square with S's belief in U.

If interpretative charity consists in maximising rationality, an interpretation of S's utterances that makes many of them false need not be all that uncharitable. Whether it uncharitable depends on *how* the interpretation makes these utterances false. If it does so by attributing lots of separate, unconnected mistakes to S, it is uncharitable. But if it makes these falsehoods consequences of a small number of mistakes, it is not nearly as uncharitable. In the case of the misguided universalist who sees tables where there are only simples arranged table-wise, we see that someone believing large numbers of falsehoods need not be utterly divorced from reality. A mereological nihilist could make sense of the universalist's beliefs, for they will recognise those beliefs as corresponding to beliefs that the nihilist themselves has about what would hold under the counterfactual (perhaps counter-possible) supposition that simples always compose. Though mistaken, the universalist's beliefs are perfectly intelligible.

Let's see how this applies to what Hirsch says. He writes:

“Now, let us imagine a community of people who “talk like perdurantists.” Shown an ordinary wooden stick they will assent to the sentence, “In front of us there are a succession of highly visible wooden objects that persist for a moment and then go out of existence.” If they are speaking E[ndurantist]-English, in which this sentence is false, they are making a mistake about what visible objects exist before their eyes. Moreover this perceptual mistake is linked to their general a priori mistake of asserting the false E-English sentence, “Any persisting object a priori necessarily consists of a succession of temporal parts.” (Note the relevance here of my general assumption that ontological disputes concern matters of a priori necessity.) Barring some extraordinarily unlikely explanation for why they would be expected to make mistakes of this sort, charity to use requires us to interpret them as speaking P-

English, in which these sentences are true.”⁴³

An endurantist interpreting the talk of this community certainly would have the option of interpreting them as speaking an English of their own, which would make all their utterances come out true. But there is also an interpretation available that, while perhaps not exactly as charitable, is still pretty charitable. That interpretation – call it the 'common language interpretation' – treats all their assertions to the effect that objects in front of them have temporal parts as false, but attributes to them only one mistake: that of thinking that “Any persisting object a priori necessarily consists of a succession of temporal parts.” That one a priori mistake explains all the other false utterances, and so the speakers of the community don't come out as wildly irrational.

There are two things the deflationist could say here. First, they might say that even if the common language interpretation is not a great deal less charitable than the one that has them speaking P-English (call this an 'alternative language interpretation') it is still, on balance, less charitable, and so not preferable. That makes it relevant for us to ask whether there are any other considerations that might rule in favour of the common language interpretation, and I will say something about that below.

Second, the deflationist might insist that the single mistake that the common language interpretation attributes to the community is a particularly egregious one, which by itself makes that interpretation quite uncharitable. Here's why they might say that: the a priori principle about temporal parts that the members of the community assent to is one that by their lights is trivially true; it is a platitude that they have learned to assent to in the process of learning their language. And we should be particularly charitable to platitudes: if an interpretation doesn't make those come out true, that interpretation is seriously flawed.

That platitudes in particular should be interpreted to come out true is certainly plausible. But rather than deciding things in the deflationist's favour, this should make us question the way Hirsch sets up his thought-experiment. The situation he

43. Hirsch (2009), p. 240.

describes is, of course, an idealisation: in actual fact there are no communities systematically disagreeing about part-hood and the like, only individual philosophers disagreeing about such things. As mentioned in section II, Hirsch has a reason for idealising in this way: he worries that there might be no respectable sense in which one can interpret the language of an individual, given that individuals count as speaking the particular language that they do in virtue of belonging to some linguistic community or other⁴⁴.

Idealisation can be very helpful and perfectly harmless, of course. But if it is to be helpful, then the conclusions we reach about the idealised situation shouldn't depend on any specific features of the idealisation that are absent, or very different, in the original situation. If that's the case, then it would be illegitimate to apply those conclusions to the original situation.

And there's the rub. Philosophers do not usually regard a priori necessary principles about part-hood as platitudes. Even if they think these principles are a priori, and are moreover strongly convinced of their truth, they still regard them as contentious and unobvious. Consequently, an interpretation of a philosopher's utterances that make such principles come out false is not exactly lacking in charity to platitudes. Rather, it is lacking in charity to controversial theoretical claims, for that is what these principles are. And while it is plausible that one ought to be charitable to platitudes, it is not nearly as plausible that one ought to be charitable to controversial theoretical claims. A far better attitude for an interpreter to take would be to always set controversial theoretical claims aside, when interpreting a community's language. After all, such claims are so often false that they don't serve as reliable evidence of meaning.

If that is right, then as far as charity is concerned, the common language and the alternative language interpretations are in a draw. Perhaps the former interpretation attributes more falsehoods overall, but these falsehoods don't seem to add up to anything in the way of irrationality.⁴⁵

Perhaps the common language theorist could even do better than the alternative

44. See Burge (1979).

45. A similar point is made by McGrath in his (2008).

languages theorist, charity-wise. Here's some evidence we haven't considered yet: ontologists take themselves to be substantively disagreeing with other philosophers about what principles about part-hood are true. They would also make utterances to that effect, if prompted. The alternative language interpretation makes these meta-dialectical utterances come out false. It makes philosophers come out as misapprehending the nature of their business. The common language interpretation makes the philosopher's meta-dialectical utterances come out true, and makes them come out as correctly apprehending what it is that they're engaged in. So it seems that the common language interpretation is charitable to meta-dialectical utterances. Should we weigh charitableness to meta-dialectical utterances heavily? Perhaps we should: would it not be irrational for someone to misunderstand what it is that they do for a living?

In response to this point a deflationist might say that the fact that such meta-dialectical utterances come out false on the alternative language interpretation is not to be weighed too heavily, because it is, after all, a controversial matter whether ontological disputes are verbal or substantial, and we have already remarked on the dubious status of a principle of charity to controversial claims. That may be right, but it's not obvious. For while those claims may be controversial among meta-ontologists, they may not be controversial among ontologists, and the latter are the people we're tasked with interpreting.

There may be more moves to be made here. I'm happy, however, if I can just show that Hirsch's considerations of charity don't clearly favour an alternative languages interpretation of ontological talk. And it seems that they don't. Hirsch's case for deflationism is not conclusive.

V. Eligibility and Apriority

Sider's view of quantifier meanings is boldly externalist: the meaning of "∃" is definitely not 'in our heads', since it is determined by whatever happens to be the most eligible candidate meaning for "∃", out there. In his (2001a), he seems to

speak as if **D**, the domain, is the meaning of the quantifier. Domains, by themselves, are not the right sort of thing to be the meaning of a quantifier – if they were, then the universal and the existential quantifier would presumably have the same meaning, which clearly isn't the case. Presumably Sider has something more subtle in mind. We can charitably assume that, whatever details of the quantifier semantics he has in mind in his (2001a), **D** plays a prominent part in it, and so it is reasonable to accord it the role of externalist meaning-determiner in his metasemantics.

His (2009a) goes into a bit more detail on the semantics, and suggests that we think of the meaning of “ \exists ” as (the nominalised counterpart of) a special property of properties (i.e. instantiation). In any case, it is the externalist character of Sider's quantifier metasemantics – combined with the metaphysical claim that the world comes ready-made with the right sort of natural feature – that guarantees that “ \exists ” means the same in the mouths of disagreeing metaphysicians, contra Hirsch.

This externalism, I'm going to argue, leads to unacceptable results. The role that eligibility plays in Sider's metasemantics for “ \exists ” opens up strange possibilities which, I'll argue, force the ontological realist to give up the plausible view that whether the usual quantificational inference rules are valid is an a priori matter.

In the abstract, and while remaining neutral on the metaphysical underpinnings, we can think of quantifier meanings as functions from properties to truth values. When it comes to specifying our metasemantics, we can then ask the general question: what does it take for some such function to be a good candidate meaning for “ \exists ”?

For some function from properties to truth values to count as a candidate meaning for “ \exists ” it must render true, to a decent extent, the logical and extra-logical claims made using quantifiers. But the best candidate meaning needn't be the one that gets the most truth out of our talk. What also counts is which candidates are the most natural and hence eligible ones. Sider allows that the overall best candidate may be one that renders false many of our ordinary existential claims: he considers the mereological nihilist quantifier meaning (according to which there are

no composite objects) a serious contender, even though it renders false most ordinary existence claims.⁴⁶

Sider believes that on this metasemantics, the meaning which ultimately gets selected as the best overall candidate meaning for “ \exists ” (at least as employed in the philosophy room) is the one which makes it true to say of all and only the things that exist, that they exist. He gets this result by means of a metaphysical hypothesis. In his (2001a) it’s the hypothesis that **D**, the collection of all the things that there are, is a highly natural thing. In his (2009a), he favours the hypothesis that instantiation is a highly natural property of properties⁴⁷. As I will try and show, if we grant either hypothesis, and also grant that eligibility can trump charity to the extent that Sider thinks it can, Sider cannot fully guarantee the result that he wants. He cannot guarantee that this particular candidate meaning will be selected as the meaning of “ \exists ”. I will first focus on the hypothesis that **D** is highly natural, and then say something about the other option. As it’ll turn out, the two face much the same problem.

Let’s grant that our world comes equipped with some highly natural properties. Some of these don’t stand a chance of being the meaning of “ \exists ”, because they make too little sense of our quantificational talk. Take for instance the property of being negatively charged. If we interpret “ \exists ” as the function that takes all the properties co-instantiated with being negatively charged to the True, we would have many cases in which existential generalisation fails (i.e. the principle that for any non-empty name a and predicate F , $\lceil F(a) \rceil \vdash \lceil \exists xF(x) \rceil$). We do however believe of quantification that it obeys this principle. In less formal terms: it should be true to say of any object whatsoever that it exists. But if we interpret ‘exists’ to mean the same as ‘is negatively charged’, that won’t be true.⁴⁸

The property of being negatively charged is a transparently bad candidate meaning for ‘exists’. But there are more challenging cases. Imagine that **D**, the class

46. In fact he thinks it’s the best candidate bar none: see his (2013).

47. Or that some nominalist substitute for instantiation is highly natural. Sider (2009a) discusses some options without definitely endorsing one specific hypothesis.

48. This informal way of putting it shouldn’t be made to bear too much weight; I don’t wish to suggest that existence is a predicate.

of all things, has a subclass **A** that includes all of the members of **D** except for one object β . Imagine, furthermore, that β is wildly dissimilar from every single object in **A**. β could be the only non-physical object in a universe of physical objects, for instance: it lacks the property of physicality (call it **G**) which all other things have.⁴⁹ I'm going to assume that one feature in virtue of which a class of things can count as natural is the objective similarity of the things that are its members.⁵⁰ In terms of that feature, then, **A** has an advantage vis-à-vis naturalness that **D** doesn't, given that **D** includes β . In virtue of including β , **D** incurs a naturalness penalty. This leaves open the possibility that **D** is overall the more natural candidate, for I am not assuming that the naturalness that Sider attributes to **D** must be due to the objective similarity of its members – Sider might have some entirely different explanation in mind, or believe that no explanation is in order.⁵¹ Still, if we assume that objective similarity is at least a contributing factor to the overall naturalness of **D**, it doesn't score as highly if it would if it didn't include β . **A** does better than **D** on that point.

Now suppose we interpreted “ \exists ” as the function that takes all the properties co-instantiated with **G** to the True. This interpretation would be almost exactly as charitable to our quantificational talk as interpreting it in terms of **D**. (Remember that the interpretation in terms of **D** could itself, according to Sider, be fairly uncharitable and yet be the best candidate meaning for “ \exists ”.) But assuming, for the sake of argument, that our language has a name for β , this interpretation will do worse in one instance, because existential generalisation will fail in the case of β , and so existential generalisation will not come out as a valid inference rule.⁵²

49. Imagine if you will a hunk of ectoplasm, or a unique divine being.

50. We must also, of course, grant to Sider the assumption that classes can be the sorts of things that have a degree of naturalness.

51. Although in his (2009b, section 9) he does consider an interpretation of the naturalness of the quantifier in terms of similarity. There it is similarity between facts, however, not objects.

52. I am assuming here that if one's quantifiers fail to range over some object α , that doesn't preclude one from having a name that refers to α . There may be theories of proper names on which names are implicitly quantificational devices (Russell's theory of descriptions, for example) and on those views this assumption might be problematic. But I'm taking it that the mainstream view is that proper names do not function in that manner. If, however, one cannot name things one cannot quantify over, that would block the argument. But the Siderian realist would then have a different problem: whether one could use “ \exists ” to quantify unrestrictedly would then be hostage to a posteriori matters (i.e. the existence of something like β).

This difference in charity between the two interpretations tells against the interpretation of the quantifier in terms of **A** and for the interpretation in terms of **D**. But charity, let us remember, was not supposed to count for everything. By increasing, in our imagined scenario, the dissimilarity between β and the other objects in **D**, we increase the relative naturalness that **A** earns in virtue of not including β , and the unnaturalness penalty that **D** gets for including it. There will be a point at which the internal homogeneity of **A** will make such a difference that the overall naturalness of **A** will outscore the overall naturalness of **D**, even if we factor in the naturalness that **D** has which was not due to the similarity of its members.

There will then also come a point at which the superior eligibility of **A** will trump the superior charity of interpreting the quantifier in terms of **D**. At that point, the interpretation of the quantifier in terms of **A** will be the best interpretation, by Sider's metasemantic principles.

We could, in principle, discover that our world is like the scenario imagined: the possibility is not a priori ruled out. If we do, we discover thereby that existential generalisation is not a valid inference rule. If this is in principle discoverable, then although existential generalisation is presumably valid, it is not so *a priori*.⁵³

Let me shore the argument up a little. The scenario I've just sketched may be thought to rely on too many questionable assumptions about the mechanics of naturalness. If so, no matter – the argument can be streamlined to do without them. Let us treat naturalness as a 'black box', and assume nothing about how and why things are as natural as they are. What we still do know is that it's supposed to be a posteriori how relatively natural things are. It is not, therefore, a priori

Another worry in the same ball-park might be that if " \exists " picks out the function that takes all the properties co-instantiated with **G** to the True, then the domain of " \exists " will be **A**. Then, one might think, one could not have a name for β because one can only name things in the domain. This is a difficult matter. The view that one can only name things 'in the domain' looks pretty appealing. After all, for them to be in the domain just is, in some sense, for them to be available to talk about. But that domain need not be the same one that the quantifier ranges over. I would suggest that, when everything goes smoothly, the naming and quantifying domains simply turn out to be the same thing. But when weird stuff happens as described in the scenario above, one might just get the result that the two come apart, and there are more things in the naming domain than in the domain of quantification.

53. This is essentially a Twin Earth argument (à la Putnam 1975). For principles of logic to be a priori, the meanings of the terms of logic must be 'untwinearthable', as Chalmers (2012) puts it. But on Sider's semantics, untwinearthability can't be guaranteed.

excluded that some candidate meaning for “ \exists ”, one that does not validate existential generalisation, is not sufficiently more natural than **D** to be the meaning of “ \exists ”. And so on Sider’s metasemantics for “ \exists ”, it is not a priori that existential generalisation is valid.

The argument makes trouble for a specific combination of views: the view that the meaning of the quantifier is determined partly by an external, worldly factor and the view that this factor is **D**, the ‘ready-made’ domain. But as I’ve already noted, Sider is not wedded to the latter view, and in later writings seems to tend more to the view that it is a highly natural and hence eligible property of properties that plays the part of the external meaning-determining factor. However, the argument still goes through, *mutatis mutandis*, for this other view. If the things we are comparing for relative naturalness are second-order properties, it remains the case that their degree of naturalness is not a priori. There will be epistemically possible scenarios, then, on which some second-order property is more natural than the one that Sider has in mind.

Let’s call the ‘intended’ property of properties “**I**”, for “instantiation”. Any property instantiated by anything has this property. Interpreting the quantifier in terms of this property would validate existential generalisation: given that a sentence of the form “ $F(a)$ ” can only be true if the property that “ F ” stands for is instantiated, there will be no sentence of that form which doesn’t entail the sentence “ F is instantiated”, or, in other words, the sentence “ $\exists xF(x)$ ”. Let’s say that some other second-order property **I*** (instantiation*) is had by almost all the properties that are instantiated, but not all. If treated as the meaning of “ \exists ”, it does less well than **I** in validating existential generalisation because some sentence “ $F(c)$ ” (where “ c ” is a name for something instantiated by a property that doesn’t have the second-order property **I***) would not entail the sentence “ $\exists xF(x)$ ” (which would then be equivalent to “ F is instantiated*”).

But **I*** may nevertheless turn out to be extremely natural, and more natural than **I**. The difference may count for more than the charity advantage (i.e. validating existential generalisation) that **I** has over **I***. If that scenario obtained, then **I*** would be the meaning of the quantifier, and existential generalisation wouldn’t be valid. It

is not a priori that this state of affairs does not obtain; hence it is not a priori that existential generalisation is valid.

So whichever of these two views one takes of the external factor which by hypothesis is determining the meaning of the quantifier, it seems one ends up having to give up that existential generalisation is an a priori valid rule of inference. Giving up the apriority of a standard principle of logic is a big bullet to bite. Speaking for myself, I would rather give up ontological realism, if that's what it takes. We see, then, that Sider's premise (d) is overpowered. As we'll see though, that doesn't at all put ontological realism out of commission.

VI. The Sider-Hirsch Dialectic

If Sider's argument for ontological realism has unacceptable consequences, would we be driven toward Hirsch's ontological deflationism? In this section, I'll argue that, for several reasons, that is not the case. If we abandon Sider's externalist metasemantics for "∃", and go along with Hirsch's, this does not mean we would thereby concede that ontological realism is false.

First, it is not clear that the ontological realist needs to provide any argument for their position. Sider and Hirsch seem to take it that meta-metaphysical claims need arguing for, while when it comes to metasemantic claims (about what fixes the meaning of "∃") it is okay to just pick one's theory and run with it. It is not clear that it's better to disagree fundamentally and intractably about metasemantics than to disagree fundamentally and intractably about meta-metaphysics. Or so I'll argue.

Second, there is no forced march from conceding that ontological debates are verbal to conceding that there are no non-language-relative facts about what there is. I'll argue that conceding the former is in fact consistent with resisting the latter.

After arguing for these two points, I'll then argue for a stronger claim, namely that we don't have to give up Sider's metasemantic style of arguing for ontological realism. I'll introduce a distinction on the level of quantifier semantics, i.e. between two notions of a quantifier meaning, that lets us propose a different metasemantics

for quantifiers that's more fit for purpose. That's for section VII, though.

VI.I. What Ought we Disagree About?

As noted before, Sider and Hirsch pursue their meta-metaphysical debate by metasemantic means: they argue about what determines quantifier meanings. This is an interesting approach, but the metasemantic claim that Sider argues from – that something in reality, a ready-made domain **D** or an eligible property of properties, acts as a 'reference magnet' for "∃" – is not the only way of defending the meta-metaphysical claim that ontological disputes are substantive.

In fact, it's useful to note that there's not much of a dialectical advantage to be had for Sider by arguing in the metasemantic way. Presumably it would appear unsubtle if Sider baldly asserted, in the face of Hirsch's talk about charity and differently quantifying languages, that there was a non-language-relative fact of the matter about what exists. But ultimately he avoids this bald assertion only through the comparably bald assertion that there is a highly natural and hence eligible candidate meaning for "∃" out there, which on his metasemantics implies the meta-metaphysical claim fairly directly. And Hirsch is under no pressure to accept this metaphysical claim.

I don't mean by this that Sider is begging the question. I think he is more charitably read as aiming simply for a dialectic stalemate with the deflationist, which would make it acceptable for him to argue from unshared premises⁵⁴. But if a stalemate is the aim, baldly asserting the meta-metaphysical claim itself won't be

54. At least his 2001b is plausibly read this way. In his 2009b he presents himself as aiming higher. He suggests a Quinean argument for believing in a distinguished quantifier meaning, writing (p. 7):

"Why believe in a distinguished quantifier-meaning? My answer is that we generally attribute distinguished meanings (meanings that carve at the joints) to the primitive expressions of our most successful theories. That is why we think that the primitive predicates of fundamental physics carve at the joints. But quantifiers occur in every successful theory that anyone has ever advanced."

This argument doesn't convince, to my mind. It's doubtful whether any one of the going theories in ontology can be called successful, compared to the others. And if they are all successful (compared, say, to theories no longer defended) one begs the question against the deflationist by assuming that they all use "∃" with the same meaning.

less effective dialectically, even if it does lack subtlety. As far as dialectical effectiveness goes, there is no compelling reason to argue for the meta-metaphysical claim that there is a fact of the matter about what exists from a prior metasemantic claim about what determines quantifier meanings.

If it also turns out, as it has, that Sider's way of arguing for ontological realism has unpalatable consequences, then there is no longer a meaningful advantage to being dialectically subtle in this way. An ontological realist ought to just state, as a basic and foundational claim, that there is a fact of the matter about what exists. This should be a claim to argue *from*, rather than something to argue toward. Taking this attitude would not at all, to my mind, be unnatural or forced. If one is willing to countenance intractable disputes anywhere in philosophy (and one should, for obvious regress reasons) then meta-metaphysics seems one of the more natural places to have them. I for one would be surprised if it turned out, in the final analysis, that metasemantic disputes really are more rock-bottom than meta-metaphysical ones.

Note that taking this attitude – call it 'flat-footed realism' – doesn't commit one to believing that ontological realists and deflationists are ultimately talking past each other (a kind of meta-deflationism). There can be a meaningful choice between the two views: it's just not going to be about finding commonly accepted premises and trying to argue from those to one or the other view. The choice would be made on holistic grounds: does deflationism or realism allow one to have the package of philosophical views that is most attractive overall?

VI.II. Verbal Debates about Real Facts

If the ontological realist goes the flat-footed route, and takes it as axiomatic that there is a fact of the matter about what exist, Hirsch's charity-based metasemantics for the quantifiers will not hold any real terrors for them. Let's suppose, as Hirsch does, that the meaning of the quantifier is determined wholly by the logical inferential role of the symbol plus existence claims, and that different sets of claims

determine different meanings. Ontological debates could then be described as verbal: it would be possible for philosophers making different claims about what there is to both be right.

That does not mean, however, that there would be no non-language-relative facts of the matter about what there is. We just wouldn't be in a position to disagree about those facts non-verbally. The metasemantics of our ontological vocabulary would be sadly ill-suited to that task. But that just says something about us and the nature of our languages, not about reality.

Here's one way to think about this. Even if ontologists of different stripes use “ \exists ” with different meanings, there would remain a certain sense in which there is a *right* meaning for the quantifier, for the purposes of doing ontology. Here's how: if there is a fact of the matter about what exists, there will be a set of existence claims that 'gets it right' in the sense that, for all and only the objects that really exist, it is correct according to this quantifier meaning to say that they exist. The pursuit of ontology, on this picture, is an attempt to simultaneously speak the right language and say the right things.

This move may seem fishy. Here's what someone might say:

“You say that a certain sense remains in which there is a right meaning for the quantifier, and hence a correct way of answering ontological questions. But you can only say that because you assume that there are ontological facts: only if there are such facts can there be a way of using the quantifier which is correct in the sense you describe. In other words, you are begging the question.”

I want to resist this line of thinking most strongly. To argue this way is to assume, again, that it is incumbent on the ontological realist to provide arguments for their meta-metaphysical claims on the basis of previously established claims about metasemantics. As I argued above, we need not conceive of the dialectic in that way. It is legitimate for the ontological realist to stake their meta-metaphysical claims first, and draw consequences about the metasemantics on the basis of them. And if we do start off by claiming that there are ontological facts, we are in a

position to describe the above derivative sense in which there are correct answers to ontological questions. If we are to assume that there are no ontological facts until we prove otherwise, then indeed we would be making a dialectical misstep. But that's not how matters stand.

So to sum up the morals of this section, Hirsch's metasemantic considerations will not gain him any traction against a properly determined realist. But all that notwithstanding, the realist might still not be very happy to have to accept Hirsch's metasemantics for the quantifier, even if this metasemantics is less obviously problematic than Sider's. Disagreeing ontologists cannot, on this view, truly accuse one another of speaking falsely, even though they are disagreeing about a substantive matter. Their disagreement would have to be described, despite initial appearances, as a disagreement about who's speaking truly in the right language and who's speaking truly in the wrong language. And while that still preserves something of their disagreement, it's arguably a bit more revisionary than we'd hoped.

But fortunately we can do better. Hirsch and Sider have some assumptions about quantifier semantics in common and if we question these, we see that there's a neater metasemantics available, on that's suitable for ontological realism.

VII. Semantics and Metasemantics

The previous section argued that giving up Sider's claim that there is a privileged meaning-candidate for "∃" does not entail an automatic win for Hirsch. But there is something more we can say. As I will explain, Sider and Hirsch seem to share certain assumptions about the semantics of quantifiers, and in Sider's case, the metasemantics that he proposes for the quantifiers is shaped partly by how he thinks about candidate quantifier meanings. As we saw, his metasemantics has some unattractive consequences.

But it turns out that we can eat our cake and have it: we can have a quantifier metasemantics which allows us to combine various appealing views. In combination

with the flat-footed realism promoted in section VI, it would (a) allow for the view that ontological debates are substantive, (b) allow us to avoid the bad consequence that the inference rules for the existential quantifier are not a priori and (c) allow us to maintain that the quantifiers are properly regarded as purely logical terms. It is this last result that matters most from the perspective of this thesis as a whole, but the other two are, of course, important too.

The trick is this: if we make some adjustments to our views on the semantics of quantifiers – if we have in mind different sort of things as our candidate quantifier meanings – then our options for the metasemantics also change.

VII.I. Local and Global Quantifier Meanings

While Hirsch and Sider disagree about the metasemantics of quantifiers, they can be said to agree broadly about the semantics of them, i.e. while they disagree about what factors determine what the meaning of “ \exists ” is, they agree about what sort of thing a candidate quantifier meaning is. Though it’s never articulated very explicitly in their exchanges, the sorts of things that Hirsch and Sider have in mind as candidate quantifier meanings can be thought of, in the abstract, as functions from properties to truth values. For both the existential and the universal quantifier, candidate meanings are function which maps some properties to the True, and all others to the False. In the case of the existential quantifier, the properties mapped to the True are ones that are supposed to be instantiated by something. In the case of the universal quantifier, the properties mapped to the True are ones that are supposed to be instantiated by all objects.

But it’s possible to give a different answer to the semantic question than the one Hirsch and Sider give. Rather than thinking of candidate meanings for “ \exists ” as functions from properties to truth values, we can think of them as functions from property-domain pairs to truth values. That is to say: the meaning of “ \exists ” only determines a truth value for a property when a particular domain of quantification is specified.

In generalised quantifier theory, this distinction between these two different options is articulated as a distinction between 'local' and 'global' quantifiers.⁵⁵ A local quantifier is a set of subsets of a domain. For a given domain **D**, the existential quantifier is the set of all the non-empty subsets of **D**. The universal quantifier of **D** is the set of the subsets of **D** that contain everything in **D** (in other words: **D** itself). A global quantifier is a function from domains to sets of subsets of those domains. So the global existential quantifier is the function that takes one from a given domain to the set of all its non-empty subsets. So for **D**, we get the existential quantifier for **D**, for some other domain **D'** we get the existential quantifier for **D'**, etc.

Note what difference these two notions of quantifier meanings make. Call the set of non-empty subsets of **D** (the local quantifier for **D**) "**EL(D)**" and call the global existential quantifier "**EG**". **EL(D)** is domain-specific: if I were to assign **EL(D)** as the meaning of my symbol " \exists ", I would thereby automatically specify that " \exists " ranges over **D**. If however the sorts of things I was assigning as the meanings of my quantifier were global quantifiers, and if I assigned **EG** to my symbol " \exists ", then I would not thereby have specified what domain " \exists " ranges over. I would only have specified that, whatever domain " \exists " ranges over, " \exists " picks out the set of non-empty subsets of that domain. So if the meanings of my quantifier symbols were local quantifiers, then by assigning a meaning to a quantifier I would not be settling quite as much as I would be, if the meanings of my quantifier symbols were global quantifiers.

There are thus two options on the level of quantifier semantics: go local (as Hirsch and Sider seem to implicitly do) or go global. Which option one takes has repercussions when it comes to the metasemantics.

55. See Westerståhl (2001). The local/global quantifier distinction is a distinction on the level of semantics. If we went for a global interpretation, then, if we wanted to, we could make this apparent in the syntax of our logic as well, by having a *binary* quantifier. In a logic with binary quantifiers, quantifiers syntactically have to come along with a specification of a domain. So instead of having expressions like " $\exists xF(x)$ ", we would have expressions like " $\exists/UxF(x)$ ", where " \exists/U " indicates that the quantifier ranges over domain **U**. If we interpret " \exists " as a binary quantifier, then the inference rules will be a little more intricate. In particular the rule of existential generalisation that played a role in section V would not work the same way. That rule said that for any non-empty name **a** and predicate **F**, ' $F(a)$ ' \vdash ' $\exists xF(x)$ '. In a binary-quantifier logic, the rule of existential generalisation would then say that for any domain **U** and any name **a** of an object in **U**, ' $F(a)$ ' \vdash ' $\exists/UxF(x)$ '.

If quantifier meanings were local, it would not be possible to settle the meaning of our symbol “ \exists ” by appealing only to the logical inferential role of the symbol. That information would not suffice to settle what domain “ \exists ” is supposed to range over: and to pin down a particular local quantifier, we would have to pin down a particular domain. To elaborate: we can be fully aware of the logical inferential role of “ \exists ” (its introduction, elimination and exchange rules) without being able to tell which properties get taken to the True by “ \exists ”. We’d have to bring in more data to settle that. A deflationist could look at the platitudinous existence claims that are made by the users of “ \exists ”. A realist could do some ontological research, and try to find out what really exists.

This would be different on a global view of quantifier meanings. To settle the meaning of “ \exists ”, we would not have to settle what domain “ \exists ” ranges over. On the global view, that's simply not part of fixing the meaning of “ \exists ”: that's a separate issue. To fix the meaning of “ \exists ”, we could just appeal to the logical inferential role of the symbol “ \exists ”.⁵⁶ On the global view, it wouldn't be the case that mereological universalists and mereological restrictivists disagree about the meaning of “ \exists ”, *even if Hirsch were right*. If Hirsch were right, then universalists and restrictivists would be having a verbal debate. But on the global view, the diagnosis wouldn't then be that the universalists use “ \exists ” in with a universalist meaning, and restrictivists use it with a restrictivist meaning. Rather, the universalists would be using “ \exists ” to range over a universalist domain, and the restrictivist would be using “ \exists ” to range over a restrictivist domain.⁵⁷

If one adopted a global view of quantifier meanings, the accompanying metasemantics would then be as follows. The facts about what “ \exists ” means are settled by how people *logically* use the symbol “ \exists ”: what formal rules they take it to

56. Of course, determining the inferential role of the symbol might itself be a difficult task. Here the Kripkenstein issues from chapter I would again rear their head. In chapter IV I say more about these issues. Another issue of relevance might be the ‘tonk’ problematic (Prior (1960-61), Belnap (1960-61)), which raises questions about how an inferential role might suffice to fix the meaning of a symbol. But that problem goes beyond the scope of this thesis.

57. The notion of a domain then has to be understood in an appropriately anti-realist manner. On a realist view of ontology, it makes no sense to speak of there being a universalist domain and a restrictivist domain, except in the sense that the latter could be a sub-set of the former, which is not what the deflationist has in mind.

be governed by. Since universalists, restrictivists, endurantists, perdurantists etc. all espouse the same formal rules for “ \exists ”, they would not count as using “ \exists ” with a different meaning. Whether they were or were not talking past each other would then still be left open. But that debate would not involve the meaning of “ \exists ”.

Why is it important that the debate not be about the meaning of “ \exists ”? The reason it matters is that “ \exists ” is part of the logical vocabulary. It is my overall aim, in this thesis, to argue that the logical vocabulary does not bring along with it any metaphysical commitments. If being an ontological realist meant believing that “ \exists ” represented a joint in reality, then wanting to be an ontological realist would mean believing that the logical vocabulary brings along with it a commitment to metaphysical structure. But fortunately, it doesn't mean that. One can be an ontological realist and still believe that “ \exists ” has no particular metaphysical import, because one can have a metasemantics for “ \exists ” that does not appeal to any metaphysical structure.⁵⁸

It might still turn out, at this point, that there is *ontological* structure that an ontological realist would need to appeal to (i.e. a privileged domain), but this doesn't play a role in fixing the meaning of the quantifier, so we have no reason to regard that structure as logico-metaphysical in nature.⁵⁹

58. This is setting aside, for the time being, the possibility that the quantifiers, *as purely logical terms*, commit us to some metaphysical structure because that structure is needed to avoid underdetermination challenges as set out in chapter I. That issue gets addressed in chapter IV and V.

59. However, suppose that one were for some reason disinclined to adopt a global conception of quantifier meanings. Then one could still incorporate something of the distinction in one's metasemantics, if not in one's semantics. Reflecting on the local/global quantifier distinction should make clear that there are two distinguishable aspects to determining the meaning of a local quantifier meaning. The information that bears on what type of set-theoretic entity the quantifier is (i.e. in the case of “ \exists ”, a set of non-empty sets) is distinct from the information that bears on which domain the quantifier is a set of subsets *of*. The former question is settled by the formal inferential role of the symbol, the latter question is settled by ontological platitudes (for Hirsch) or ontological facts (for Sider). One might then also distinguish, respectively, two aspects to a local quantifier meaning: call the former the *logical* aspect (settled by logical input) and the latter the *ontological* aspect (settled by ontological input). One could then make the following claim: *insofar as* “ \exists ” is a logical symbol (i.e. as far as the logical aspect of its meaning goes) “ \exists ” brings no commitment to structure with it. All the metaphysical commitments (to an ontological joint in reality or the like) attach to the ontological aspect of its meaning. And that claim would be what is minimally required, for my purposes.

VII.II. Domain Assignments

Suppose an ontological realist went along with this proposal of ‘going global’. How should they then continue their debate with the deflationist? How exactly could they avoid the problem articulated in section V?

One *prima facie* option an ontological realist like Sider has is to again appeal to an externalist hypothesis of eligibility. They could claim that there is a privileged domain **D** containing everything, which one's quantifiers range over whenever one opens one's mouth in the philosophy room. They could say that, when we assign domains to our quantifiers, it matters not only what general ontological claims one makes, but also which domains are particularly eligible assignments. The ontological realist could then contend that **D**, which by hypothesis contains everything, is in fact particularly eligible, and so one's quantifiers range over it whether one likes it or not. (And ontological deflationists could of course refuse to go along with that story, and the debate could continue.)

If the realist were to do that, however, they would still run afoul of something analogous to the problem described in section V. There is still the epistemic possibility that some set of objects **D*** happens to score better vis-à-vis eligibility than **D**. If that happened, then **D*** would be the domain of the existential quantifier. And then, again if we allow names for things that the quantifier fails to quantify over, we end up with possible failures of existential generalisation. Existential generalisation would be a priori valid.⁶⁰

But fortunately, when we have the ‘flat-footed’ ontological realism argued for in section VI, we do not have to go this route. Our story about domain assignments can be similarly flat-footed. Flat-footed realists, from the get-go, accord themselves the right to talk about ‘what there really is’. They do not go through the metasemantics to secure the result that there is a non-language-relative fact of the matter about what there is. Having assumed this privilege, they can use it when

60. And, again, if one did not concede that there could be names for things the quantifier failed to range over, there would be the milder but still unattractive consequence that it would not be a priori that the quantifier, when unrestricted, ranged over all of the things that exist.

they get around to specifying their metasemantics. What is the domain that the quantifier, when unrestricted, ranges over? It is the class of the things that exist. Why is D^* not the domain assigned to the quantifier? Because, by hypothesis, there is something that isn't in D^* , and that rules out D^* as the domain assigned to the quantifier.

If that sounds a bit crude, that's because it is. What I'm suggesting that it's okay, in this instance, to be crude. However, there are more subtle ways of specifying domain assignments available to the flat-footed realists. In the appendix to this chapter I go into some more detail, as a way of dealing with the issue of Skolem-style indeterminacy about unrestricted quantification.

VIII. Conclusion

In this chapter we have discussed another for logico-metaphysical structure: the argument from ontological realism. I've argued that ontological realism does give us some reason for making metaphysical claims about the world: it give us a reason to claim that there's a privileged domain of quantification. But it doesn't give us a reason to believe in any structure that's distinctively logical in nature. We don't need to believe that the quantifiers themselves (*sans* domains) carve at any joints.

This chapter and the previous one form a unit. They constitute the 'negative' part of this thesis. They have been concerned with setting out some arguments that purport to show that the world has structure corresponding to our logical constants. I have offered reasons to doubt these arguments, and in that way I've already argued a little for my own, more minimalist view: that we should not take the fundamentality of logic to our theorising to imply, of itself, that logic is fundamental to reality.

In my view, logic is fundamental in a sense, but that sense is not a metaphysical one. Logic is not baked into reality. The rest of this thesis sets out a view of the logical vocabulary with exactly that upshot. It does not suppose that there is logico-metaphysical structure, but it can still explain why logical terms crop up in our

thought, talk and theorising, and why the logical vocabulary isn't subject to indeterminacy of the Kripkenstein variety.

IX. Appendix: Unrestricted Quantification

Vann McGee (harking back to Putnam) presents us with the following worry about indeterminacy in unrestricted quantification:

The bothersome worry, it seems to me, is not that our domain of quantification is always assuredly restricted but that the domain is never assuredly unrestricted. The origin of this worry is Skolem's paradox, and it finds its sharpest expression in Hilary Putnam's 'Models and Reality'. Suppose, for *reductio ad absurdum*, that our usage picks out a unique intended model for our language, and that this model has an all-inclusive domain. The Löwenheim–Skolem theorem tells us that there is a countable set S such that we get an elementary submodel of the intended model when we restrict the domain to S . There isn't anything in our thoughts and practices in virtue of which the so-called intended model fits our intentions in using the language better than the countable submodel, so that there isn't anything that makes unrestricted quantification, rather than quantification over S , the intended meaning of the quantifiers.⁶¹

McGee discusses some responses to this problem, eventually offering his own, but it looks like an ontological realist of Sider's variety has a fairly easy time with it. They can say that unless the quantifier is explicitly restricted by means of some predicate or by a context, or perhaps expanded by some fictionalist operator, there is an interpretation for it out there that's more eligible than S and renders the quantifier fully unrestricted (and also not under-restricted). Barring any funny business on our

61. McGee 2006, p. 185, Putnam 1983.

part, the meaning of “ \exists ” will be the naturally eligible one, and in this meaning, the quantifier really does range over all and only the things that exist.⁶²

This Skolemite issue doesn't arise for an ontological deflationist, or at least is not as worrying. According to the deflationist, the question of what there is for the unrestricted quantifier to range over and the question of what our thought and talk requires the quantifier to range over are one and the same question. Insofar as our thought and talk leaves open how we should interpret the quantifier, there is no ontological fact of the matter either. As far as ontology goes, nature does not draw sharper boundaries than language does – or so the deflationist holds. Because the deflationist is not as worried by the issue raised by McGee, it would be dialectically important for an ontological realist – one that doesn't take Sider's particular line – to have an adequate response to it, or else that could be held against ontological realism. Fortunately, ontological realism brings resources along with it that we can appeal to in responding to the problem.

In a moment I will present my favoured solution to the Skolemite worry, but first I want to discuss another suggestion that may seem quite compelling. Couldn't a non-Siderian ontological realist deploy something like the Siderian strategy outlined just above? To be sure, an ontological realist who adopted the quantifier semantics advocated in the preceding section wouldn't be in a position to make that exact response. But something quite similar is available to them.

On the global view of quantification, for unrestricted quantification to take place is just for the quantifier to range over the biggest one available, rather than some subset of that biggest domain. For an ontological realist, the biggest domain will just be 'the' domain **D**, the one containing all and only the things that really exist. Let's suppose that we're in a context where nothing is going on that would restrict the domain. What would it take in that situation for the quantifier to take **D** as an argument, rather than **S**? If fitting with our thoughts and practices doesn't do the job, then maybe we should just hypothesise, à la Sider, that **D** is highly natural and thus a more eligible candidate than **S**. But, *pace* Sider, we wouldn't be proposing

62. I'm not actually attributing this response to Sider; I put it forward as a response that one might find appealing if one had the resources of Sider's ontological realism at one's disposal.

that it's an eligible *meaning* for the quantifier. Rather, it would just be an eligible domain-assignment for the quantifier⁶³.

I would think that if the Siderian response outlined above works, then this response would work just as well. Unfortunately, the Siderian response doesn't really work as well as one would hope. An appeal to eligibility may well rule out **S** as a meaning-candidate for the quantifier – so far so good – but as we saw in section V, Sider cannot a priori rule out the above-mentioned **A** as a candidate. The realist is left hostage to scenarios in which eligibility makes the quantifier's meaning be a restricted one, and while the specifically Skolemite worry is avoided, unrestricted quantification still isn't guaranteed. The strategy which we adapted from the Siderian one, which talks about eligible domain-assignments rather than eligible quantifier-meanings, would face an analogous problem.

Luckily there is something else the realist can say which does better. By hypothesis, interpreting our quantifier in terms of **S** makes as much sense of our thoughts and practices as **D** does. If our thoughts and practices had been different, **S** might not have gotten the job done as well as **D**, but in that case, by parity of Skolemite reasoning, there would have been another set, **S'**, which would have done the job as well as **D**. Note, however, the difference between the following two claims:

- 1) Whatever our thoughts and practices, there is a proper subset of the domain that makes as much sense of them as the domain does.
- 2) There is a proper subset of the domain that, whatever our thoughts and practices, makes as much sense of them as the domain does.

While 1 is true (as a result of the Löwenheim–Skolem theorem) 2 is not. That matters. Given that **S** is a proper subset of the domain, there is at least one thing that isn't in **S**. Suppose our thoughts and practices are such that **D** and **S** make equal sense of them. Take an object **α** that's in **D** and not in **S**. If our thoughts and

63. The notion of eligibility wasn't designed to apply to things like domain-assignments; I assume for the sake of argument that it's extendable to them.

practices were different in such a way that α were referred to and talked about,⁶⁴ then \mathbf{S} would not have been as adequate to our thoughts and practices as \mathbf{D} .

What ontological realists can and should say is that the best candidate domain-assignment for the quantifier is not just one that makes sense of all our thoughts and practices, but one that makes sense of all our *counterfactual* thoughts and practices that are precisely as follows: they are like our thoughts and practices, except that each of them brings up one of the objects that's out there, but isn't talked about in our thoughts and practices. Under those conditions, \mathbf{D} does better than any \mathbf{S} -like set, because for any \mathbf{S} -like set, there is a counterfactual set of usage facts such that \mathbf{D} is adequate to those facts, but that particular \mathbf{S} -like set is not. This, we should say, is what makes \mathbf{D} the domain that the quantifier really ranges over.⁶⁵

It seems that this shows us the way out of the woods as regards unrestricted quantification. Not quite, though: there is a further issue. Having taken a global view of the quantifier, I've so far interpreted McGee's worry as being about a possible indeterminacy in which class of things the quantifier takes as its domain. But one might think that Skolem-style indeterminacy can crop up in other ways. For we may worry that the quantifier itself (seen, in the binary/global manner, as a function from domains and properties to truth values) fails to range 'properly' over the whole of \mathbf{D} , even if it does determinately range over \mathbf{D} and not over \mathbf{S} . The quantifier could be interpreted to 'see' only the \mathbf{S} -sized fragment of \mathbf{D} , even though it ranges over the whole of \mathbf{D} (i.e. it would take to the False properties that are only instantiated by those things in its domain that are in the complement of \mathbf{S}). For the same Skolemite reasons, interpreting the quantifier as being (so to say) *S-myopic* would be just as adequate to our thoughts and practices as interpreting it in the straightforward way.⁶⁶

I am not fully convinced that this further worry is real. It might just be that the myopic quantifier could be ruled out because they didn't behave in a properly 'logical' way: for one thing, these sorts of quantifiers wouldn't seem to obey

64. In short, if our thoughts and practices had been in some way such as to make it unavoidable to involve α in their interpretation.

65. Though different in the details, this solution is akin to McGee's own preferred solution.

66. On the unary/local view, this worry wouldn't come apart from the initial Skolemite worry.

existential generalisation, and that might rule them out. But if we concede, for the sake of argument, that they are real candidate quantifier meaning, then we see that this worry can be answered in a way that's similar to how we addressed the initial worry. For note the difference between:

- 3) Whatever our thoughts and practices, there will be a myopic quantifier meaning that makes as much sense of them as the straightforward quantifier meaning.
- 4) There is a myopic quantifier meaning that, whatever our thoughts and practices, makes as much sense of them as the straightforward quantifier meaning.

Once again, the former is true but the latter is not. So realists should demand that an adequate interpretation of the quantifier make sense of more than just our actual thoughts and practices. It ought to make sense of counterfactual thoughts and practices that are like our thoughts and practices, except for mentioning some object in the quantifier's domain that isn't mentioned in our actual thoughts and practices. This will exclude all myopic quantifier meanings as inferior candidate quantifier meanings, in the same way that all S-like sets are excluded as inferior domain-assignments.

Note that these strategies rely straightforwardly on ontological realism: because we start out assuming that there is a real, non-language-relative fact of the matter about what things there for quantifiers to potentially range over, we can help ourselves to these objects when we're specifying the adequacy conditions for interpretations of quantifications. From some perspectives this may seem cheap. But from the perspective of ontological realism, it's perfectly alright.

Some might also think I'm making it too easy on myself in another way: by supposing that the language which I'm using to state my metasemantic proposal is such that its quantifiers do determinately range over everything. I compare **D** and **S**, and say that **D** contains everything and **S** doesn't. But how do I know that "everything" in the previous sentence doesn't range over **S** rather than **D**? If I don't, am I not just begging the question?

I don't think I'm begging the question here. The meta-language in which I'm stating the adequacy conditions for interpretations of the quantifiers in the object language has quantifiers in it, and so there should be a story about why *these* quantifiers can range over everything. (I've not distinguished very assiduously between object and meta-language in this section, but I could have done so, at some cost to readability.) However, such a story is available: I've told such a story already about the object language, after all. So the recipe would be: just repeat that story in a meta-meta-language, and if you're wondering about the quantifiers of *that* language, repeat it again at a yet higher level.

Obviously, an infinite regress would result from this. But that's only a problem if the regress is vicious, and I don't think it is. For the task at hand is not that of *giving* meaning to the quantifiers of our object language: they already have meanings, and we're just trying to explain in virtue of what facts they have their meanings. We're articulating the meta-semantics that's already in force, not designing a semantics that's yet to be implemented. And our meta-language in this case is just English, a natural language whose words similarly already have meanings. If we were giving meanings to words which don't yet have meanings, and we try to do this by using words which also don't yet have meanings, we wouldn't get anywhere: a truly vicious regress would arise. Explaining how it is that words that do have meanings have their meanings, by using words that also already have meanings, is potentially regressive, in the sense the questions one answers at one level have analogues at a higher level. But it's not viciously regressive, for at any level at which one might want to pose the questions, a complete answer is available.

Chapter III: Logical Expressivism

0. Introduction	78
0.I. Some Preliminaries and Caveats	81
I. Propositional Attitudes, Simple and Complex	84
I.I. Rejection	84
I.I.II. A Spectrum of Acceptance and Rejection	87
I.II. Complex Propositional Attitudes	89
I.III. Why Complex Attitudes Are Propositional Attitudes	91
I.IV. Generating Greater Complexity	93
II. Sequentese	95
II.I. The Symbol “▶”	96
II.II. Some Connectives Replaced	97
II.III. Embedded Sequents	100
III. Pinning Down Sequentese	102
III.I. Syntax of Sequentese	103
III.II. Translating Sequentese	104
III.II.I. Translating Sequentese: Sequentese to English	105
III.II.II. Translating Sequentese: from PL to Sequentese	108
III.II.III. Translating Sequentese: from Sequentese to PL	111
III.II.IV. Translating Sequentese: Recoverability	113
IV. Reasoning in Sequentese	115
IV.I. Motivating the Inference Rules	118
IV.II. Motivating the Structural Rules	124
IV.III. Embedding and Extraction Rules for Sequentese	129
IV.IV. Is Sequentese Classical?	132
IV.IV.I. Classical Derivability Implies Sequentese Derivability	133
IV.IV.II. Sequentese Derivability Implies Classical Derivability	136
V. Conclusion	141

0. Introduction

In previous chapters, we've looked at one view of the logical vocabulary and its metaphysical significance: Sider's logical realism. I've argued against that view. But it has the virtue of explaining why we have the logical vocabulary: logic is out there in the world, and we need a way to talk about it. So if we're going to abandon that view, there had better be some other way of explaining the presence of the logical vocabulary that we can put in its place. That's what this chapter is for.

So why do we have logic? More particularly, why do we have terms like 'and', 'or', 'if', 'not', and 'every' in our vocabulary? It's not just English that has such terms, of course: most natural languages, as far as we know, have terms that are more-or-less equivalent to these, or at least can play very similar roles. And then of course there are formal languages like the language of first order predicate logic, and various extensions of it. A recognisably similar set of terms crops up in all of these.

In earlier parts of this thesis we've heard one story explaining the presence of these terms in our languages. That's the story in terms of logico-metaphysical structure. On this story, the logical terms signify some aspect of the world. They don't necessarily stand for *things* in the world – they needn't be names – but they reflect something real, and the fact that they reflect something real explains why we do well to have these terms in our languages. If we didn't, we'd be missing something that's out there.

But that story has the drawback – or so I've presented it – of giving us some substantial metaphysical commitments. It may not be entirely clear what these commitments amount to, since it's not entirely clear what logico-metaphysical structure is, but it's clearly something over and above the various other kinds of structures that we may or may not be committed to – physical structures, mathematical ones, mereological ones and so on. As a good Ockhamite, I'd like to do my bit to help us all avoid multiplying structures beyond necessity.

That's what this chapter is about. It presents a view of the logical vocabulary, a theory of meaning to be precise, that explains why we have the logical terms and what they do for us. This story is very unlike the one that Sider gives us, because it denies that the logical terms signify anything (though not that they are *significant*, in a general, non-referential sense). It denies that there's anything in reality that the logical vocabulary helps us get at. And since it denies this, the explanation for why we have these terms cannot be that there are some joints in reality that want carving.

The story instead will be that the logical vocabulary plays a *merely expressive* role. We need it because without it, we couldn't express certain attitudes to propositions that we need to be able to express if we are going to have constructive exchanges of views, agree about things, disagree about things, make arguments, convince each other, and so forth. Given that this kind of communication is important, I take it, the logical terms do an important job for us.

An important term here is *merely expressive*. Just to say simply that the terms of logic play an expressive role would not do anything to set them apart from non-logical terms. All terms are expressive devices. But the way that a term like, say, 'apple' helps us express something is by referring to something; by standing in some semantic relation to something in the world. The term can contribute that which it stands for to the content of a sentence, and thereby helps us express attitudes to propositions that are about apples. This is not the way in which the logical terms help us express attitudes. The logical terms do not contribute anything to the content of what's expressed: they help indicate what sort of attitude towards the content is being expressed.

For the most straightforward example, take the term 'not' (or 'it is not the case that'). To view this term as a merely expressive device is to claim that by adding it to a sentence we don't change anything concerning what that sentence is about – the content remains the same. What we change is what attitude we express concerning that content. If I say 'apples are fruits' I express a belief that apples are fruits or, to put it in slightly different terms, I express my acceptance of the proposition that apples are fruits. If I say 'apples are not fruits' or 'it is not the case

that apples are fruits' I don't express my acceptance of a different proposition, i.e. the proposition that apples aren't fruits, but my rejection of the original proposition, that apples are fruits.

This example is exceedingly simple, misleadingly so even, and we'll see that things get a lot more complicated when we move to other terms of logic, or even just to negations that do not take outer scope. But it does make clear the basic idea that this chapter is going to elaborate on: that logical terms indicate what attitude one is expressing, rather than indicating anything about the content of the attitude. To spell out that idea, I'm going to introduce a set of propositional attitudes that's rather more expansive than the usual set (Section I). I'm going to introduce the notion of *complex propositional attitudes*, attitudes which involve more than one atomic proposition and which have a rich internal structure.

As part of spelling out that proposal, I'll present a formal language which I call *Sequentese*, which is distantly based on a multiple-conclusion sequent language (Section II). For this language I'll give a syntax (Section III.I), a systematic semantics in terms of complex propositional attitudes (III.II.I), and a scheme for translating between it and the language of classical first-order logic (PL) (Sections III.II.II – III.II.IV). Sequentese will serve as a stepping stone for providing PL with an expressivist semantics in terms of complex propositional attitudes.

But we'll do more than that. On the basis of just a few claims about the rational relations that complex propositional attitudes stand in to each other – to be precise, claims about which attitudes cannot rationally be combined – we can motivate a set of inference rules for Sequentese, turning it into a propositional calculus (Section IV.I & IV.II.). It turns out that it is a fully classical calculus which has the same consequence relation as classical propositional logic: the premises and conclusion of any valid classical PL argument, when rendered into Sequentese, deliver an argument that's valid in the Sequentese calculus, and vice versa (Section IV.III). That will help to show that the ideology of complex propositional attitudes, implemented by means of the Sequentese formalism, gives a viable and non-revisionary account of the logical relations that sentences of PL stand in, an account

which does not need to appeal to the properties of the logical connectives in those sentences.

After setting out how all of that works, we can go on to explain why it is that the logical constants have determinate meanings. But that will take us into chapter IV.

0.I. Some Preliminaries and Caveats

Before going on, I should say a little more about exactly what I mean by *the* logical vocabulary or *the* terms of logic. The short answer is: I mean the logical constants. But that may not help much, for there is a big and lively debate on what it takes for a term to be a logical constant, and little in the way of a consensus. I am going to try and skirt that debate for now, although I hope to comment on it in a later chapter. But I think I can skirt this debate, for while there is no consensus on what the conditions for logical constant-hood are, there is something of a consensus on which terms count as such. That is to say: there is a set of terms such that one would *prima facie* expect a theory of logical constant-hood to rule these terms in, and others out.

The terms in question are the logical connectives (\vee , \rightarrow , \wedge , \neg , \leftrightarrow , and others in the same class of interdefinables, like the Sheffer stroke) and the universal and existential quantifiers (\forall , \exists). If a theory of the logical constants failed to rule any of these in, it would presumably be regarded as a non-starter. At the same time, it's not inconceivable that a well-motivated theory of the logical constants could rule out some of these (the quantifiers, let's say) without thereby ruling itself out. The notion of logicity seems sufficiently theoretical that an acceptable theory of it could be revisionary to some extent.

Indeed, the set of purported logical constants just cited appears at least a little heterogeneous. Strikingly, introducing quantifiers into one's language makes one's semantics considerably more complex, but once the semantic notions that we need to deal with quantification are on the table, then expanding one's language to include modal operators – if they have the Kripkean semantics that we standardly

attribute to them – isn't quite as big a jump in complexity. Of course the matter is extremely arguable, but a case could be made that the quantifiers group together more naturally with the modal operators than they do with the connectives.⁶⁷ And if the modal operators are definitely not logical constants, as some would insist, then what are we to say about the quantifiers?

At the end of this thesis, I'll say more about considerations like these, although it won't be my concern to argue for any particular view of logicity. It does turn out, in any case, that the most tractable way to set out the expressivist view is by first tackling propositional logic (i.e. the connectives) and only then moving on to the quantifiers. So in this chapter, I'll restrict my attention to propositional logic. Quantifiers will keep until Chapter V. By doing things that way we also make the proposal somewhat more modular: if for some reason one is less convinced, at the end of the day, of the logical expressivist treatment of quantification than of the logical expressivist treatment of the connectives, one could take the latter and leave the former.

One further point before we get started. I have waved away, for the moment, some difficult questions about logical constant-hood. For our immediate purposes, to be a logical constant is just to be one of the symbols usually so classified. Our 'expressivist' account of the constants will count as successful if it handles those symbols successfully. But one might think that still doesn't give us a clear target yet. There may be one negation *symbol*, one conditional *symbol*, etc., but according to different logics, the behaviour of these symbols is very different. If the expressivist account makes " \rightarrow " come out as behaving classically, is that a good result? In this respect, the conditions of success are underspecified.

One way of dealing with this issue would involve engaging in the debate between classical logicians and proponents of the various non-classical logics and coming to some conclusion about what the 'right' behaviour of the logical constants would be.

67. As I say, it's extremely arguable. On the one hand, we can appeal to the fact that a Kripkean semantics for the modal operators has some structural analogies with the standard semantics for quantifiers, in the sense that the modal operators get treated sort-of like quantifiers over worlds. On the other hand, modal operators are syntactically unlike quantifiers and more like connectives, in that quantifiers bind terms and modal operators and connectives don't. But the point is simply that a philosophical debate could be had about these matters.

We could then take that behaviour as our target, i.e. a successful expressivist semantics would have to make the constants exhibit that behaviour. For a number of reasons, that's not what I'll do. First off, it wouldn't be feasible: making any significant intervention in the 'which is the one true logic' debate would be simply too big a project to undertake as a part of this one. Second, whether we should be looking for the 'one true logic' in the first place is an open theoretical question in its own right. Various philosophers – 'logical pluralists' – have argued that there can be a meaningful sense in which multiple logics are correct.⁶⁸ It seems that we would also have to settle that debate, if we want to do things properly.

Instead, I'll take a more pragmatic approach. I'm going to nominate classical logic as the *pro tanto* target: the expressivist semantics ought to make the logical constants exhibit classical behaviour. The reason for picking classical logic is that it's a simple, straightforward and well-understood logic that has many adherents. Justifiably or not, it enjoys the status of being the default logic. Developing a logical expressivism that can cope with a classical take on the logical constants will hopefully constitute a kind of 'proof of concept' of the expressivist project, i.e. it will show that it can work in at least one prominent case. Then one might go on to see whether different versions of the expressivist semantics might be available which are suited to other logics. If there are, we could eventually develop a more general and comprehensive expressivist proposal, one that would show how to give an expressivist semantics for any given logic.⁶⁹ That's how things would hopefully pan out. But in this thesis, I'll restrict myself to giving an expressivism that copes with classical first-order predicate logic and hope that, in showing that such a view is workable, I will have shown something significant.

Let me add one final caveat. There is another view to be had which also can be called logical expressivism, though it might also be called – more accurately – *meta-logical* expressivism. This is the view that in making meta-logical claims, i.e. claims about derivability, consistency, logical truth, and the like, we don't make factual

68. The *locus classicus* of this fairly recent debate is Beall & Restall (2005).

69. It may be the case, however, that a fully general approach is simply not on the cards. The proposal that I'm going to develop in this chapter looks like it would be quite hard to adapt to paraconsistent logics, for instance, and paracomplete logics also present some *prima facie* problems. But hope springs eternal.

claims but express attitudes to these propositions (or possibly to their forms). A view of this sort is defended by Michael Resnik in his (1996) and (1999). This is not the view I'm defending. Though my sort of logical expressivism does have consequences for how meta-logical claims are to be understood (they are to be understood as ultimately claims about the rationality of having certain (combinations of) propositional attitudes) it does not imply that these claims are non-factual. Of course, if one held a further view to the effect that claims about rationality are to be understood in non-factual terms (a view defended by Alan Gibbard in his (1990)) then my sort of logical expressivism would imply Resnik's sort of logical expressivism. But I will not be concerning myself with the general status of rationality, and so will leave it open how logical expressivism (in my sense) relates to meta-logical expressivism.

I. Propositional Attitudes, Simple and Complex

I.I. Rejection

The logical expressivism that I'll defend appeals to an unusually rich repertoire of propositional attitudes. To start off, it takes as primitive not only the propositional attitude of *acceptance* but also the attitude of *rejection*. By rejection, I understand a way of being opinionated about a proposition that is similar, but opposite in valence to acceptance. Rejection is not the mere absence of acceptance: it is a substantive state of mind whose presence or absence is not constitutively tied to the presence or absence of acceptance.⁷⁰ (I should note that this way of using the term "rejection" is fairly standard, but not universal: some would use the term in such a way that suspensions of judgement count as rejections. I am not using the term in that way: suspensions of judgement, on my use, count as situations in which one neither accepts nor rejects a proposition.)

70. So it's at least conceptually possible that someone would accept and reject the same thing, or do neither. That said, these things may of course be ruled out as a matter of rationality, i.e. it may never be a good idea to accept and reject the same thing, even if it is possible.

Why take rejection as primitive? The most obvious reason for this is one of explanatory direction: on the standard way of treating rejection, due to Frege, the rejection of a proposition P amounts to nothing more than the acceptance of its negation. But since I'll want to explain negation in terms of rejection, I can't analyse rejection in terms of negation, on pain of circularity. There are, however, good general reasons for taking rejection as primitive that do not simply drop out of the aims of the current project.⁷¹

There is a growing literature in the philosophy of logic on the various things one can do when one has recourse to both the notion of acceptance and the notion of rejection (or, on the level of speech acts, assertion and denial). There are technical issues that prove more tractable once the ideology has been enriched in this way, such as Carnap's 'categoricity problem'⁷², and there are also more strictly philosophical applications⁷³. In particular, the appeal to rejection is unavoidable if one wants to give an account of the pragmatics of disagreement that fits with paraconsistent logic. If one believes that it is sometimes rational to believe a proposition and its negation, for instance in the case of the Liar, one can't hold that disagreement consists in one person believing a proposition and another person believing its negation. If you and I were both adherents of paraconsistent logic, and both had a dialethic solution to the liar, we should count as agreeing, not disagreeing. The solution is to say that disagreement consists rather in one person accepting a proposition and another rejecting it; for the paraconsistent logician, accordingly, rejecting a proposition is something distinct from accepting its negation.⁷⁴ Similarly, an intuitionist has reason to take rejection seriously, for it will not do for their purposes to equate rejecting a proposition with accepting its negation. After all, there will be cases in which, according to the intuitionist, the rational thing to do is to reject a proposition and also reject its negation: cases where, for some proposition P , neither P nor $\neg P$ is warranted.⁷⁵

71. Frege (1919)

72. Smiley (1996), Rumfitt (2000). For a different technical application, see Bendall (1979).

73. See for example Price (1990 and unpublished) and Restall (2005).

74. Priest (2005), Restall (ibid.).

75. Restall (ibid.).

These considerations about paraconsistent and intuitionist logic will not, perhaps, seem to have much force for someone with a prior commitment to classical logic (or some other non-classical logic). But even if one's commitment is to classical logic, one might well want to be in a position to make sense of the things that paraconsistent logicians and intuitionist logicians say. A significant worry is that, given the different behaviour that logicians of different stripes ascribe to negation, a debate between them *about* negation may start to look merely verbal, each participant being right in describing the behaviour of the negation symbol in the way they do, given the notion they've chosen to let it pick out. Having recourse to a notion of rejection that's not defined in terms of negation would allow paraconsistent logicians, intuitionists and classical logicians to spell out their disagreements in a not merely verbal way. They can argue and disagree substantively about which acceptance-rejection behaviours should be considered kosher, and having had that debate, they could then go on to decide which notion of negation is the right one: it should be one which legitimises only acceptance-rejection behaviours that are kosher. So a primitive notion of rejection is useful for classical philosophers of logic that want to engage with non-classicists, even if one could do without it for the purposes of studying classical logic in isolation.⁷⁶

For our purposes, helping ourselves to a primitive notion of rejection will be a first step to giving an expressive account of the logical connectives. The most immediately available application is in explaining what negations do when they take outer scope: they turn the sentence that they attach to into a vehicle for the expression of a rejection, specifically a rejection of whatever it is the rest of the sentence says. As far as our aims go, that's something, but it's not much. That's only one connective, and it's not much good if there's nothing we can say about negations that don't take outer scope. We're going to need to appeal to more than just acceptance and rejection. A lot more, in fact: I'm going to argue that for every possible sentence-form of first-order propositional logic, there's a propositional attitude. But fortunately, we won't have to add each and every one of those

76. Ibid.

propositional attitudes as a new primitive of our theory. We can generate them systematically out of a small number of basic building blocks.

I.I.II. A Spectrum of Acceptance and Rejection

We first need to narrow down the notion of rejection a little more, though. Often acceptance is equated with (all-out) belief and rejection is correspondingly equated with disbelief. But this is not the only option. One could opt for a broader notion of acceptance that include belief, but also cognate attitudes like supposition and imagining, and a corresponding broader notion of rejection. Attitudes like belief, supposition and imagining are related in that they all involve taking a stand in favour of a proposition, and in that sense they are all different from attitudes like wondering and considering. They are different from each other in that they are normed by different things.

Belief is normed by evidence (there must be a certain amount of evidence in favour, and not too much against). Supposing is sometimes *negatively* normed by evidence – i.e. things against which one has evidence or knows to be false are not valid candidates for supposition. That would be the case with what is sometimes called supposing-as-actual, as opposed to supposing-as-counterfactual. The latter might be thought to be subject to a weaker negative norm: one can only suppose as counterfactual things which one does not know to be impossible. Imagining might be governed by a spectrum of norms, depending on context. For instance, if one is imagining that some propositions hold as part of the process of writing, say, a historical novel, one's imaginings would be normed by consistency with known historical fact. If one is engaging in some science fiction, one's imaginings might be normed by consistency with known science. And if one is engaging in some surrealism, none of the above might apply. The corresponding rejection-attitudes would be governed by norms which are the converse of the norms that govern the acceptance-attitudes. To disbelieve something, one ought to have sufficient evidence against it, and not too much evidence in favour of it. To 'counter-suppose'

(for lack of a better word) something as actual, it ought to be something that's not known to be true. And so on for the other types.

Different attitudes on the acceptance spectrum thus differ in the norms that govern them, and the same goes for the different attitudes on the rejection spectrum. There are some norms, however, that apply to all these attitudes. For instance, one's beliefs, suppositions and imaginings must all be consistent in the sense that it's incoherent, rationally, to believe and disbelieve the same thing (insofar as it is even psychologically possible, of course), to suppose and counter-suppose the same thing, or to imagine and counter-imagine the same thing. Such combinations of attitudes are ruled out from the get-go, so to say: regardless of what the facts are, what one knows to be true or false, or what kind of evidence one has. It is a norm that's internal to the attitudes. To accept and reject the same thing is not so much a cognitive misfire as it is a cognitive malfunction.

This consistency norm isn't the only norm that I take to be internal to the (broadly construed) attitudes of acceptance and rejection. In section IV I will suggest some other norms of this sort, which relate to the complex propositional attitudes that I've yet to introduce (see the next few sections). The importance of these sorts of norms is that they are of particular *logical* significance, or so I'll argue. The normativity of logic, I will suggest, is ultimately just the normativity of rationality. But that doesn't mean that every norm of rationality is a norm of logic: rather, the norms that are internal to the attitudes of acceptance and rejection, and the complex attitudes that involve them, are the particularly logical norms of rationality.

Because this is the aim, it's more illuminating to adopt the more inclusive notions of acceptance and rejection, which include attitudes beyond belief and disbelief. If we were looking just at belief and disbelief, then we might get the impression that all the norms that govern belief and disbelief in particular are norms that govern acceptance and rejection. And if we took all of those norms – including evidential norms – as 'logical' norms, we'd arrive at a set of inference rules that we wouldn't easily recognise as 'purely logical'. But it will be much easier to explain how and why this is the case when we've got more of the logical expressivist machinery set

up. I will return to this topic in section IV. For now, it should just be noted that I'm using acceptance and rejection in the broader senses just indicated. That being said, it will do no harm, throughout most of this chapter (until we reach section IV) to think of acceptance and rejection as just belief and disbelief, if one finds that this makes it easier to work with the notions.

I.II. Complex Propositional Attitudes

Towards a given simple proposition A, there are two basic stances to take: accept it, or reject it. That gives one, in the abstract, four options for what to do. One could just accept A, just reject it, do both, or do neither. Once more than one proposition is in play, the choices multiply. With two propositions A and B in play, one can have sixteen distributions of attitudes, towards three propositions, sixty-four, and so on.

Call a full set of attitudes towards a set of propositions a *pattern*. A pattern can be good or bad for a variety of different reasons. Take the minimal pattern that just consists in me accepting the one proposition A. This can be a good idea, if A is true, and a bad one, if A is false. Or take the plan that consists in me accepting A while rejecting B. Suppose that A in this case is [it's raining] and B is [the streets are wet]. Since the streets are generally wet when it's raining, this would be a bad pattern: when accepting A is a good idea (i.e. when A is true) rejecting B is a bad idea (because it would be false). Or take the pattern that consists in me accepting A and rejecting A: this is a bad idea simply because it's a paradigmatic failure of rationality to accept and reject the same thing.

Complex propositional attitudes are attitudes to patterns. The way that I will spell out the story, they are always one of two different types of attitudes to patterns. On the one hand, there is the attitude of being unwilling to instantiate a particular pattern. If you will: *ruling out* a pattern. Suppose I rule out the pattern where I accept A and reject B: that means that I will let my attitudes to A, and my attitudes to B, be constrained in a certain way. If, as it happens, I accept A and I've not formed an opinion on B, then my options for going ahead and forming an opinion

on B are limited to accepting it. Or if I already accept A and reject B, I will revise my opinions so that I stop accepting A, or stop rejecting B.

On the other hand, there is the attitude of wanting to instantiate a certain pattern. I will call that *ruling in* a pattern. Suppose I rule in the pattern of accepting A and rejecting B. This, like ruling out, also constrains my attitudes to A and B, though in a much more direct way – there is only one way in which one can make one's attitudes to A and B fit with the ruling-in of this pattern, namely by accepting A and rejecting B.

Ruling in and ruling out stand in a similar relation of opposition to each other as acceptance and rejection do. Neither consists simply in the absence of the other, and so it is possible, in the abstract that is, to take any one of four stances to a pattern. One might (i) just rule it in; (ii) just rule it out; (iii) do neither or (iv) do both. Ruling in and ruling out are both substantive ways of being opinionated about something, and in that sense they are both opposed to the stance of neither ruling in nor ruling out a pattern of attitudes.

It is plausible that there are situations in which one ought not (in an internal sense, so by one's own lights) rule in or rule out a pattern: to wit situations in which one has no evidence concerning the merits of the pattern. (Of course, depending on assumptions about what makes patterns good or bad, it may be the case that in a more objective sense of 'ought', it is always the case, for a given pattern, either that one ought to rule it in or that one ought to rule it out.) By contrast, taking the stance of ruling in and ruling out the same pattern, like accepting and rejecting the same proposition, looks to be straightforwardly irrational from both an internal or an external perspective. And just as it's plausible to think that in the case of acceptance and rejection, the irrationality of doing both to the same proposition is what makes these attitudes be essentially opposed to one another, one can see the irrationality of ruling in and ruling out the same pattern as constitutive of the opposition between these attitudes. I will say more about the connections between these attitudes and norms of rationality below, in section IV.

Complex propositional attitudes, as I've construed them, can be fruitfully compared to plans. Suppose I know that you have a plan to get some ice cream if

the thermometer goes above 25° Celsius. That's not the same as knowing whether you're getting some ice cream. But given some extra information, it will allow me to figure out whether you're getting some ice cream. Your plan leaves open what your precise actions will be, but it nevertheless constrains them. Similarly, knowing that you rule out the pattern where you accept A and reject B doesn't tell me whether you accept A or whether you reject B. But if I find out that you accept A, I can deduce that you don't reject B. And if I find you that you reject B, I can deduce that you don't accept A. Complex propositional attitudes constrain attitudes like plans constrain actions. Of course, a plan can leave nothing open at all, and in those cases your plans will tell me what you will do without me having to bring any other information to bear on the matter, except of course information about your ability to do what you plan. Similarly attitudes of ruling in leave nothing open: if I rule in accepting A, accepting B and rejecting C (say) then I know what your attitudes will be to those propositions, on the assumption that no funny business (like epistemic *akrasia*⁷⁷) goes on.

In another way complex propositional attitudes are also different from plans, of course. Actions, and intentions to act, are generally under the control of the will. Propositional attitudes, simple and complex, generally are not. I don't generally just decide to accept a proposition, or to rule out a pattern. Usually, such things seem rather to be imposed on me, by the world or by some norms of rationality. When I perceive that A, accepting A is generally the only option that's open to me, psychologically. And the reason that I don't accept and reject one and the same proposition is not that I've made a prior decision not to: it's that I feel I should not, or cannot even do so. So nothing I've said about complex attitudes should be taken to express a commitment to any form of voluntarism about acceptance and rejection attitudes.

I.III. Why Complex Attitudes Are Propositional Attitudes

77. I'm not assuming that epistemic *akrasia* is the sort of thing that ever happens, but if it does, I know what your attitudes are only on the assumption that it didn't happen in your case.

What makes complex propositional attitudes, as I've here described them, *propositional* attitudes? There are two aspects to that question. First, it seems that as I use the term, complex propositional attitudes are not attitudes to propositions at all, but attitudes to patterns, i.e. sets of propositional attitudes. So do they deserve to be called propositional attitudes? Second, how do complex propositional attitudes, again as I use the term, relate to attitudes to complex propositions, i.e. attitudes to propositions which have some logical structure? I'll answer these questions in turn.

With regard to the first question, I want to suggest that complex propositional attitudes are two-headed beasts. On the one hand, they are attitudes to (sets of) attitudes. That's how I've introduced them. But they are *simultaneously* attitudes to basic propositions. That may sound a little strange, but there is no real reason why they can't be both at the same time. To illustrate, suppose again that I rule out accepting A while rejecting B⁷⁸. This is an attitude to a pattern, and patterns are sets of attitudes. But simultaneously, it's an attitude to A and B: a complicated, multi-place attitude. To rule out a pattern of attitudes to A and B is genuinely a way of being opinionated about A and B. Since ruling out accepting A while rejecting B is consistent with a bunch of sets of attitudes to A and B (accepting both, rejecting both, etc.) it doesn't amount to a fully determinate way of being opinionated about A and B. But it is nevertheless a way of being opinionated.

With regard to the second question, I want to suggest that there are, in the final analysis, no attitudes to logically complex propositions. There aren't even any logically complex propositions to have attitudes towards. The complex attitudes I've introduced are meant to replace attitudes to complex propositions in our philosophical scheme of things. In those instances where we would be inclined to say that a subject accepts a proposition with some logical complexity going on, I

78. In the next few sections, I will often be using phrases like 'ruling out accepting A and rejecting B'. This is meant to be short for 'ruling out the pattern of accepting A and rejecting B'. The latter phrasing makes explicit that ruling out is not distributive: by ruling out the pattern of accepting A and rejecting B I do not individually rule out accepting A and rule out rejecting B. What I'm ruling out is the combination of attitudes. Also, both of the above phrases are subject to a scope ambiguity: they can be read as "ruling out (accepting A and rejecting B)" or as "(ruling out accepting A) and (rejecting B)". Unless explicitly stated otherwise, the former reading is always intended, where the "ruling out" takes wide scope over the "and".

suggest that they have a complex attitude to some simple propositions. Suppose a subject says “If it snows, the trains will be delayed”. One way to explain this is as the expression of the acceptance of a proposition of conditional structure. But it can also be explained, and I submit that we should so explain it, as the expression of a complex attitude to two propositions: that it is raining and that the trains will be delayed. The subject is committing themselves to accepting that the trains will be delayed if they accept that it snows, and rejecting that it snows if they reject that the trains will be delayed. In other words, they rule out accepting the former while rejecting the latter.

Is this revisionary? I don’t think it is. Or, if it is, then not in an objectionable way. It is novel, for sure. But the explanation of the subject’s utterance as the expression of the acceptance of a complex proposition is given by a theory: a popular, uncontroversial theory, but a theory still. The datum by itself – the utterance of a complex sentence – does not commit us to it. And to explain this utterance as the expression of a complex propositional attitude does not, I believe, do any violence to it as a datum.

In the course of this chapter, I will try and show that there are no cases in which we can’t replace the ‘simple attitudes to complex propositions’ story with a ‘complex attitudes to simple propositions’ story, if the logical complexity is that of the sentential connectives. We do not need to suppose that there are propositions of conditional, negated, disjunctive or conjunctive structure – all that work can be done by complex attitudes. In chapter five, I will extend the account to quantificational structure as well, but since that brings in a lot of extra complications, I will keep it out of the picture in this chapter.

I.IV. Generating Greater Complexity

It will not be clear at this point how it can be guaranteed that for any purported complex proposition, there is a complex attitude that can do its explanatory work. We can guarantee this by allowing patterns of attitudes to be embedded in other

patterns: in effect, to allow patterns of attitudes to figure other patterns in much the same way that atomic propositions do, even though they are not themselves propositions.

As defined, a pattern is a set of attitudes to propositions. But let's change that a little: let a pattern be a set of attitudes to propositions and patterns. That may sound a little circular, but the notion a pattern can be recursively defined. Let something be a pattern if it is one of the following:

(1) A level 1 pattern: a set of attitudes (of acceptance and rejection) to atomic propositions;

(2) A level 2 pattern: a set of (a) attitudes of acceptance and rejection to atomic propositions and/or (b) attitudes of ruling-in and ruling-out to level 1 patterns;

(3) A level 3 pattern: a set of (a) attitudes of acceptance and rejection to atomic propositions and/or (b) attitudes of ruling-in and ruling-out to level <3 patterns;

...

...

...

(N) A level n pattern: a set of (a) attitudes of acceptance and rejection to atomic propositions and/or (b) attitudes of ruling-in and ruling-out to level <n patterns.

Take one of the two attitudes to patterns that we've introduced above: that of ruling it out, i.e. not wanting one's propositional attitudes to instantiate it. Let those attitudes of refusal crop up at the same level that we've had attitudes of acceptance and rejection cropping up, and suddenly we can make sense of a lot more seeming logical complexity. Suppose someone says "If it's snowing, or there's a storm, the trains will be delayed". Standardly, we'd take this to express a conditional with a disjunctive antecedent: something of the form " $(A \vee B) \rightarrow C$ ". But we can explain this as the expression of a complex propositional attitude as well.

It's the attitude where we rule out the following: ruling out rejecting both A and B, while we're also rejecting C.

Above, I argued that attitudes to patterns are *ipso facto* propositional attitudes: by refusing a certain pattern of propositional attitudes, one takes an attitude towards the propositions involved in that pattern. By ruling out certain combinations of acceptances and rejections, one narrows down one's opinions about the propositions in question, even if one doesn't fully determine them. This claim – that what I've called propositional attitudes are genuinely attitudes to proposition – is not compromised by allowing attitudes to patterns to figure in patterns themselves. For these more complex, layered patterns still bottom out in attitudes just to propositions. Take the example above: the attitude of ruling out (ruling out denying both A and B) while (rejecting C). We can still legitimately view this as a propositional attitude whose objects are A, B and C. The layering introduces more complexity into the attitude, but it doesn't change what it's ultimately about.

II. Sequentese

The conceptual building blocks of the logical expressivism I want to defend have now all been introduced: the propositional attitudes of acceptance and rejection; the notion of a complex propositional attitude as an attitude to a pattern of propositional attitudes; the notion of embedded patterns. But it needs to be shown that these notions give us enough power and flexibility to motivate the claim that all attitudes to logically complex propositions are in fact complex attitudes to sets of simple propositions. We can only do that persuasively by tackling things in a more systematic, formal manner. In the next sections, I'm going to set out a formal language in which we can generate, for any sentence of propositional logic, a representation for a complex attitude which it expresses. It'll be called Sequentese.

This language is inspired, as far as its syntax goes, on multiple-conclusion sequent calculus.⁷⁹ But unlike standard sequent calculus, there are not going to be any connectives in this language. And also unlike standard sequent calculus, it's going to be possible, in this language, to embed sequents in sequents to create more complex, layered sequents. This will allow us to represent attitudes to layered patterns of attitudes. In the next three subsections, I'll introduce the bits and bobs of this language informally, to show what the rough idea is. Then I'll go on to specify the syntax of the language in full, and give a method for translating between this language and standard propositional logic.

II.I. The Symbol "►"

I will introduce Sequentese in two stages: first a simple version of the language which does not have any possibility of embedding, and then a fuller version which does allow embedding.

In Simple Sequentese, we express propositional attitudes by means of multiple-conclusion sequents of the general form:

$$A_1, A_2, A_3, \dots \blacktriangleright B_1, B_2, B_3, \dots$$

Such a sequent expresses an attitude of ruling out: what is ruled out is accepting all of the propositions on the left-hand-side while rejecting all of the propositions on the right-hand-side. If it happens that one does accept all of the propositions/rule out all of the sequents on the left-hand-side, this means one cannot reject all of the propositions/rule in all of the sequents on the right-hand-side. But if one does reject all of the propositions on the right-hand-side then not everything on the left-hand-side can be accepted. One person's *modus ponens* is another's *modus tollens*. As is the case with multiple-conclusion sequent calculus, there can be any number of names of propositions on either side of the symbol – including zero.⁸⁰

79. The possibility of using multiple-conclusion sequent calculus to dispense with the logical connectives is anticipated in Akiba (1996).

80. I owe the idea of interpreting a multiple-conclusion sequent in terms of acceptance and

The above sequent bears a close resemblance to a sequent of multiple conclusion sequent calculus, save that the normal turnstile “ \vdash ” has been switched out for the little black triangle (the ‘pointer’, as I will call it). The reason for this switch is to avoid confusion: sequents with the turnstile usually indicate that the propositions on the left-hand-side *logically imply* the ones on the right-hand-side. For the purposes of developing Sequentese, we don’t want anything that strong. The pointer does express something that’s structurally quite like a consequence relation. But the relation is weaker. As noted above, one could adopt the complex attitude of ruling out accepting that it is snowing while rejecting that the trains are delayed. That might be expressed with the pointer as follows:

That it is raining \blacktriangleright that the trains will be delayed

But there is of course no relation of logical consequence between the two propositions. There is, however, what we might call a relation of *material* consequence between these two propositions. Given the way things actually are (as opposed to how they have to be, by dint of logic), a correct pointer sequent codifies a truth-preserving inference. The relation of material consequence, though weaker than logical consequence, is structurally analogous to the stronger relation. I will motivate this in detail below.

II.II. Some Connectives Replaced

With the pointer and its syntax introduced, I will now explain how to express a few relatively simple attitudes in Simple Sequentese. I will explain how simple assertions and denials are expressed: my approach is to treat those as limit cases of the expression of complex attitudes. I do that for the purpose of keeping the Sequentese simple: there may be purposes for which it would be worthwhile to enrich the language with some more symbols, to wit unary operators which directly express the acceptance or rejection of a proposition. But for our purposes, we’re not going to need symbols like that.

rejection to Restall (2005).

I will also explain how some slightly more complex attitudes, involving multiple propositions, are expressed. This will illustrate, at a basic level, how the connective-less language of Simple Sequentese can mimic the behaviour of some of the connectives of standard propositional logic. All this will give a first taste of how the Sequentese language, once fully described, can be viewed as a connective-free alternative to the standard language of propositional logic (henceforth “PL”).

To start, take the following sequent:

► B

This expresses that one rules out rejecting B (by expressing the ruling out of the pattern which consists simply in the rejection of B). That, of course, isn't the same as expressing the acceptance of B, but it is close to it in the sense that, if one rules out rejecting B, the only way to expand one's set of opinions is to accept B. A sequent of this form, then, allows one to do something quite like asserting a proposition. There's two ways in which we might shore that up a little. One is to suggest that unless there's something funny going on, as a matter of psychology people will accept the propositions the rejection of which they rule out. This does not seem too implausible to me. Another is to simply stipulate that in Sequentese, by ruling out a rejection one conversationally implies an acceptance.

Now take the following sequent:

A ►

This expresses that one rules out accepting A (by expressing the ruling out of the pattern which consists simply in the acceptance of A). A sequent of this form allows one to signal the denial of a proposition: if one cannot accept a proposition, the only attitude that one can adopt toward it is rejection. Expressing a rejection is, on the face of it, akin to expressing the acceptance of a negated proposition. Indeed, on some views (in particular Frege's) of rejection, that is exactly what the expression of a rejection does. It will emerge, further below, that the pointer, in a situation like this, does indeed behave somewhat like a negation operator of classical PL.

Now for a slightly more complex case:

A ► B

This expresses that one rules out accepting A and rejecting B. This means that accepting A would leave accepting B as one's only option, and rejecting B would leave rejecting A as one's only option. So this expresses an attitude of ruling out a pattern (accept A, reject B) but *ipso facto* it expresses a propositional attitude towards A and B. Uttering a sequent like this would lay one open to both modus ponens and modus tollens arguments, for if one could be persuaded to accept A, one could thereby be persuaded to accept B, and if one could be persuaded to reject B, one could thereby be persuaded to reject A. This should make the above sequent look a lot like a classical PL conditional: " $P \rightarrow Q$ ". And that's no accident: later in this chapter, it will become clear that the pointer does indeed behave like a conditional in a situation like this.

There's also:

► A, B

This expresses that one rules out rejecting both A and B. If one does reject one, then the only way left in which one might extend one's opinions is to accept the other. That should make this sequent look a lot like a classical PL disjunction: " $A \vee B$ ". And that, again, is the idea: the pointer does indeed behave like a disjunction in this situation.

All this is to show that Simple Sequentese gives one the ability to express some things that in the language of classical first-order logic we would express with, respectively, " A ", " $\neg A$ ", " $A \rightarrow B$ " and " $A \vee B$ ". There are, however, a lot of things that one wouldn't be able to express just by using the pointer, flanked by some names of propositions. For instance, there is no simple way of expressing what a simple conjunctive sentence " $A \wedge B$ " expresses. And most sentences of standard propositional logic that have connectives in the scope of other connectives can't be translated just by arranging names of atomic propositions around a pointer.⁸¹

81. Although some can: " $A \rightarrow (B \wedge C)$ ", for instance, could be translated as " $A \triangleright B1, B2$ ". And a sequent like " $A1, A2 \triangleright$ ", which expresses the refusal to accept both A1 and A2, could be expressed straightforwardly in the language of propositional logic only if it had a NAND-connective. But these are isolated cases.

The reason that Simple Sequentese can't express these things is that it has no way to express layered patterns of attitudes. But we can remedy that by making it possible to embed sequents within sequents. So let's add that option.

II.III. Embedded Sequents

In this section, I'll introduce informally the idea of embedded sequents, and explain how it is supposed to increase the expressive power of Sequentese. Further down, the full syntax of Sequentese will be pinned down exactly.

It will be allowed, in Sequentese, to take a sequent, put some brackets around it, and then use it in the same way as one would use the name of an atomic proposition. In other words, we can have well-formed sequents like these:

$A, B \triangleright C, (D \triangleright E)$

$A, (B \triangleright C) \triangleright D, E$

$A, (B \triangleright C) \triangleright D, (E \triangleright F)$

As we've so far explained it, the location of a propositional name on the left- or right-hand-side of a pointer determines which attitude is assigned to that proposition in the pattern that the sequent expresses the ruling-out of. So the sequent " $A \triangleright B$ " expresses something about the *acceptance* of A and the *rejection* of B: namely that one doesn't want to have the combination of these attitudes. The reason that it's about the acceptance of A and not the rejection of A is simply that A is on the left-hand-side of the pointer, and not on the right-hand-side. The same applies, *mutatis mutandis*, to B.

Similarly, it matters whether we put a bracketed sequent on the left- or right-hand-side of a pointer. Different sides of the pointer are going to have different 'valences', so to say. In the case of propositions, these valences are explained in terms of two different attitudes to propositions, acceptance and rejection. To extend the use of the pointer elegantly to embedded sequents, we need to bring in

the other attitude to patterns that we've introduced: ruling in. Just as to rule out a pattern is to not want to instantiate it, to rule it in is to want to instantiate it. As noted, ruling in stands to ruling out as acceptance stands to rejection: just as acceptance and rejection are two diametrically opposed ways of being opinionated about a proposition, neither of which consists merely in the absence of the other, likewise ruling out and ruling in are independent but opposed attitudes to patterns of attitudes to propositions.

I can explain now how I mean embedded sequents on either side of the pointer to be read. First, take a right-hand-side embedding like this:

$$A \blacktriangleright (B \blacktriangleright C)$$

As every sequent does, this expresses the ruling-out of a pattern. To be precise, it would express the ruling-out of the pattern which consists in (i) accepting A and (ii) ruling in the pattern of accepting B and refusing C. In other words, if one were to accept A, then one would have to not rule in the pattern on the right-hand-side. Just as, with regard to single propositions, accepting and rejecting are the only two ways of being opinionated about them, with regard to patterns ruling in and ruling out are the only two ways of being opinionated about them. So if one were to accept A, the only way to extend one's opinions about the pattern of accepting B while endorsing C is to rule out this pattern. Of course, it also goes the other way around. If one ruled in the pattern of accepting B while rejecting C (for instance by actually accepting B and rejecting C) then one would be committed to rejecting A.

Now take a left-hand-side embedding like the following:

$$(A \blacktriangleright B) \blacktriangleright C$$

Again, it expresses the ruling-out of a pattern, as any sequent does. Here, the pattern consists in (i) ruling out accepting A while rejecting B and (ii) rejecting C. Suppose one did rule out accepting A while rejecting B. Then this would commit one to not rejecting C. Suppose, however, that one did reject C. Then this sequent would commit one to ruling in the pattern of accepting A while rejecting B (i.e. one would be committed to accepting A and to rejecting B).

I hope it's clear that there's an analogy between what happens to a proposition when its name appears on the left-hand-side of a pointer, and what happens to a sequent when it appears there, and likewise that there's an analogy between what happens to a proposition when its name appears on the right-hand-side of a pointer, and what happens to a sequent when it appears there. If a name of a proposition appears on the left-hand-side, then that proposition is not to be accepted unless what is on the right-hand-side is accepted. If a sequent appears there, its pattern is not to be ruled out unless what is on the right-hand-side is accepted (or ruled out, if it's a pattern). If a name of a proposition appears on the right-hand-side, then that proposition is not to be rejected unless what is on the left-hand-side is rejected. If a sequent appears there, its pattern is not to be ruled in unless what is on the left-hand-side is rejected (or ruled in, if it is a pattern).

Embedding is also allowed within sequents that are themselves embedded, as in the following rather complex sequent:

$$A, (B, C \blacktriangleright (D \blacktriangleright E, F)) \blacktriangleright G, H$$

Deeper embedding does not, however, introduce any new complications. In the above sequent, the ruling-out of a pattern is expressed. That pattern consists in: (i) accepting A; (ii) ruling out accepting B and C whilst ruling in the pattern of accepting D while rejecting E and F; (iii) rejecting G and H. That's obviously a mouthful, but it should be clear how it that reading was arrived at. If a sequent appears on the left-hand-side of some pointer (no matter how deeply embedded) it expresses a pattern to be ruled out only if what is on the right-hand-side of that pointer is accepted/ruled out, and to be ruled in if what is on the right-hand-side of that pointer is rejected/ruled in. If a sequent appears on the right-hand-side of some pointer (no matter how deeply embedded) it expresses a pattern to be ruled out if what is on the left-hand-side of that pointer is accepted/ruled out, and to be ruled in only if what is on the left-hand-side of that pointer is rejected/ruled in.

III. Pinning Down Sequentese

Now that we've seen roughly what Sequentese looks like, it'll be good to state exactly how the language works syntactically, how its expressions are to be rendered into English, and how this language relates to the language of standard first-order propositional logic.

III.I. Syntax of Sequentese

Let's define what a well-formed sequent of Sequentese is. First I'll define the preliminary notion of a *level 0 sequent* or *proto-sequent*. Something is a proto-sequent iff it is either:

- (1) The name of a proposition ($P_1, P_2, P_3, \dots, Q_1, Q_2, Q_3, \dots$);
- (2) The name of a set of propositions (Γ, Δ, E, \dots).

With this notion defined, we can define what a sequent of Sequentese is. Something is a sequent of Sequentese iff it is one of the following:

- (1) A level 1 sequent: the result of writing the pointer symbol " \blacktriangleright " flanked on either side by zero or more level 0 sequents separated from each other by commas;
- (2) A level 2 sequent: the result of writing the pointer symbol " \blacktriangleright " flanked on either side by zero or more level <2 sequents (surrounded by brackets) separated from each other by commas;
- (3) A level 3 sequent: the result of writing the pointer symbol " \blacktriangleright " flanked on either side by zero or more level <3 sequents (surrounded by brackets) separated from each other by commas;

...

...

...

(N) A level n sequent: the result of writing the pointer symbol “►” flanked on either side by zero or more level $<n$ sequents (surrounded by brackets) separated from each other by commas;

In Sequentese, as defined here, proto-sequents and sequents play the same syntactic role: they appear on the flanks of pointers. But they aren't representing the same sort of thing, because sequents do not stand for (sets of) propositions, unlike proto-sequents. They do not stand for anything, strictly speaking, meaningful though they are. They express propositional attitudes, and the objects of these attitudes are all the atomic propositions represented by the proto-sequents that appear in the sequent.

Another way of making the same point is to emphasise that the pointer is not a connective, even if it mimics some of the connectives in certain ways. It does not take some propositions and produce a new proposition: it produces an expression for an attitude towards those propositions. Although the Sequentese formalism treats sequents and proto-sequents as syntactically on a par with each other, they are most definitely not on a par with each other semantically.

III.II. Translating Sequentese

Sequents of Sequentese express propositional attitudes, of lesser or greater complexity. Of course, this is not something I've argued for: it's something I've stipulated, in defining up the language. What needs to be argued is that ordinary language expressions involving logical connectives (“and”, “or” and the like) and complex expressions of propositional logic express propositional attitudes of lesser and greater complexity. What this mainly requires us to show is that expressions of propositional logic can be given plausible and complete translations into Sequentese. I take it for granted that we can translate ordinary language expressions involving logical connectives into the language of propositional logic, and so it would suffice to show that the latter can be rendered into Sequentese.

But showing how to translate expressions of propositional logic into Sequentese, and claiming that these in turn express propositional attitudes, is not enough. We also need to provide the means for taking a sequent of Sequentese and figuring out what the attitude is that it expresses. In other words, we need a systematic interpretation for Sequentese. One way one might do this is by giving a formal semantics for the language. But this, I think, would not be helpful. On the one hand, since Sequentese is ultimately supposed to help us give us a new semantics for the logical connectives in expressive terms, it wouldn't be useful to apply standard-issue modal-theoretic semantics to this task of interpreting Sequentese. After all, we already interpret the logical connectives in that way, so we wouldn't have made any progress.

On the other hand, it doesn't seem useful to give an entirely new formal semantics for Sequentese. The proffered means for understanding the unfamiliar Sequentese language would then be an equally unfamiliar formal semantic notation which we would then have to explain in turn. What I'll do instead is much more basic, but hopefully more useful. I'll provide a manual for translating sequents of Sequentese into English expressions describing propositional attitudes. Those expressions will be somewhat convoluted and stilted, but I think they would still help to make the Sequentese formalism intelligible in a language which does not need any explaining in turn.

III.II.I. Translating Sequentese: Sequentese to English

It's important to note the difference between *expressing* an attitude and *reporting that one has* an attitude. Take the following sequent:

A ► B

It may be tempting to translate this as follows into English:

"I rule out the pattern of accepting A while rejecting B"

But this English sentence, while arguably capturing everything in the sequent, also adds something that's not really in the sequent: the reference to the person having the attitude. This English sentence is in part *about* the person having the attitude, whereas the original sequent is only about A and B. In the light of this, I think it would be best to render sequents into English not as sentences, but as infinitive verb-phrases: names for the attitudes in question. So for the above sequent, the translation would be:

“Ruling out accepting A whilst rejecting B”

That is the type of translation that I'll give a manual for. Here goes. To translate a sequent into English, the recipe is as follows. One starts one's translation by writing down 'Ruling out ...'. One then starts on the left-hand-side of the sequent. For any name of a proposition, one writes 'accepting [that name]'. For any name of a set of propositions, one writes 'accepting the members of [that name]'. For any embedded sequent, simply skip it for now: maybe insert some brackets '(...)' leaving them unfilled for the moment. Insert 'and's and commas where necessary to preserve grammaticality.

Once one has gone through everything on the left-hand-side in the above manner, one goes to the right-hand-side. First, write a comma and then write 'whilst'. Then, for any name of a proposition, one writes 'rejecting [that name]'. For any name of a set of propositions, one writes 'rejecting the members of [that name]'. For any embedded sequent, simply put some placeholder brackets '(...)' again. And again, insert commas and 'and's where necessary to preserve grammaticality.

Having gone through the right-hand-side this way, one turns to the embedded sequents one has skipped, if there are any. For any such embedded sequents on the left-hand-side, one first writes 'ruling out ...' and then one simply applies the instructions just above to whatever things flank that sequent. For clarity's sake, it may help to put some brackets around the whole resulting phrase. For any embedded sequents on the left-hand-side, one first writes 'ruling in ...' and then, again, one applies the above instructions to translate whatever flanks that sequent, enclosing the whole thing in brackets. If there are any embedded sequents in those

embedded sequents one has just translated, apply the same recipe to those. Rinse and repeat until nothing remains untranslated.

In this manner, one should arrive at something which is a grammatical verb-phrase of English. One can add commas to taste. Take the following sequent:

$$\Gamma, P, (T \blacktriangleright U) \blacktriangleright Q, (R \blacktriangleright S, (V \blacktriangleright W)), \Delta$$

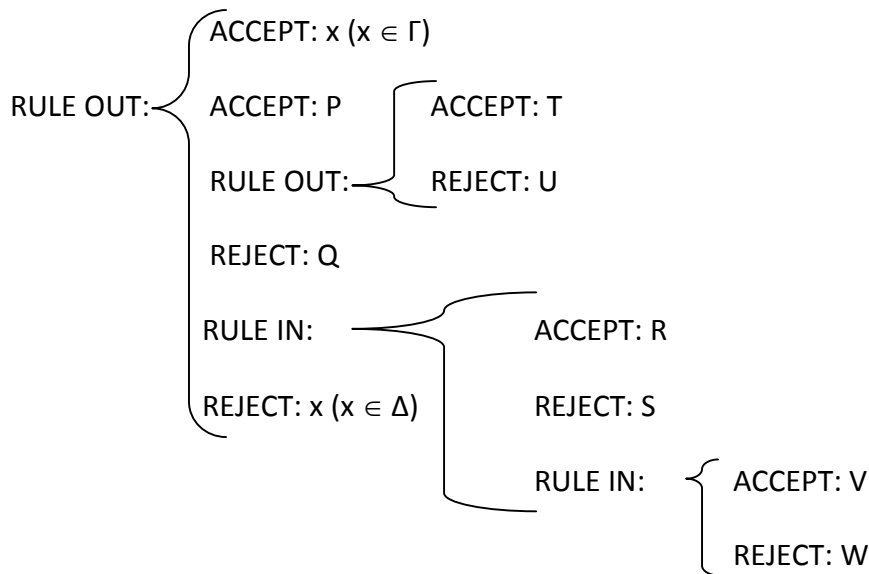
Here's the English translation one would arrive at, with a judicious sprinkling of commas and 'and's:

“Ruling out accepting the members of Γ , accepting P and (ruling out accepting T whilst rejecting U) whilst rejecting Q, (ruling in accepting R whilst rejecting S and (ruling in accepting V whilst rejecting W)), and rejecting the members of Δ .”

Now obviously this is a mouthful, and not very pretty, but it is a parse-able English phrase. The use of brackets makes it more understandable when read than spoken, but the limits of understandable spoken English are quickly reached, so not much could be done about that. There would be ways of making it less unwieldy by making it slightly less English, like for example the following:

“RULE OUT: accepting the members of Γ , accepting P, (RULE OUT: accepting T and rejecting U), rejecting Q, (RULE IN: accepting R, rejecting S and (RULE IN: accepting V and rejecting W)), and rejecting the members of Δ .”

One might do something like that, if one felt that the advantage of brevity outweighed the advantage of English grammaticality. One could go even further, but one quickly reaches the point where one might as well draw a diagram instead:



Some may find such diagrams helpful. But there will come a point where such representations stand in need of a translation manual themselves, and at that point they would no longer be helpful. I will stick with translations into relatively plain English.

III.II.II. Translating Sequentese: PL to Sequentese

Sequentese is intended to be equivalent in expressive power to the language of first order propositional logic (PL henceforth). This, of course, needs to be shown. In the following, I'll give translation manuals for translating Sequentese expressions into sentences of PL and vice versa, and a 'recoverability' result to the effect that applying both manuals in succession would give one back the expression that one started out with.

Two notes before we start. First, to keep complications to a minimum, I'll use a relatively sparse version of PL with only \vee , \rightarrow and \neg as its connectives. Since versions of PL with more connectives can be translated easily into this sparse version of PL, I take it this is not a significant limitation. Second, I will be limiting myself to providing Sequentese translations for PL sentences that are 'downwardly' finite in complexity: sentences of PL that are built up by taking some atomic

sentences, sticking those together with connectives, sticking *those* together with connectives and so on. Similarly, I will provide PL translations only for Sequentese sequents that are downwardly finite in complexity.

I'll start with going from PL to Sequentese. If the sentence of PL has as its main connective a disjunction, like so:

$$P \vee Q$$

Then it will be translated as follows:

$$\blacktriangleright P, Q$$

This expresses, according to the Sequentese-to-English translation manual given above, that one rules out rejecting both P and Q. This, then, is also what I take the original sentence of PL to be expressing. If the sentence of PL has as its main connective a conditional, like so:

$$P \rightarrow Q$$

Then it will be translated as follows:

$$P \blacktriangleright Q$$

This expresses, according to the Sequentese-to-English translation manual given above, that one rules out accepting P while rejecting Q. This, then, is also what I take the original sentence of PL to be expressing. If the sentence of PL has a negation as its main connective, like so:

$$\neg P$$

Then it will be translated as follows:

$$P \blacktriangleright$$

This expresses, according to the Sequentese-to-English translation manual, that one rules out accepting P. That is what I take the original sentence of PL to be expressing.

The sentence of PL might of course have not logical complexity at all, like so:

$$P$$

If that is the case, then it'll be translated as follows:

► P

If the original PL sentence had more than the single connective, then by applying this recipe only to the main connective one will initially get something that's neither an expression of PL nor of Sequentese. Suppose the sentence one starts with is the following:

$\neg P \rightarrow Q$

Then by translating the conditional, one gets:

$\neg P \blacktriangleright Q$

To get a proper Sequentese expression, one either needs to not start at the main connective, or now translate the PL sub-formula of the above pseudo-sequent into Sequentese. We need the former procedure, since we're assuming only that the sentences of PL are downwardly finite, i.e. that we *can* start at the bottom level. So if we start with:

$\neg P \rightarrow Q$

We need to translate " $\neg P$ " into Sequentese first, before tackling the conditional. By doing this, we end up with:

$P \blacktriangleright \rightarrow Q$

But in order not to end up with nonsense, we'd better add some brackets at this stage too. So:

$(P \blacktriangleright) \rightarrow Q$

And now we can translate the conditional, and end up with:

$(P \blacktriangleright) \blacktriangleright Q$

And this is a well-formed sequent of Sequentese.

The procedure should be clear by now: one first translates the minimum-complexity sub-formulae of the PL sentence, bracketing those translations, and then moves to the next layer of connectives, until one arrives at the main

connective, at which point one should get a well-formed sequent of Sequentese. If the sentence of PL is infinitely complex (upwardly) then one would of course never reach the completed translation. But still, one would know how to get there, and that's the important thing.

After translating a PL sentence into Sequentese, one could then use the Sequentese-to-English translation manual from the previous section to obtain a canonical English reading for the original PL sentence.

III.II.III. Translating Sequentese: Sequentese to PL

The reverse procedure is fairly straightforward as well, though it's made much easier by adding a conjunction symbol to PL. I'll assume we're using a version of PL that has one. To tackle the degenerate cases first, take a sequent with only an atomic sentence letter or set-name on the left-hand-side:

A ►

This can be translated as follows:

¬A

Take a sequent with an atomic sentence or set-name only on its right-hand-side:

► A

This will be translated as:

A

There may be nothing on the right-hand-side, but more than one thing on the left-hand side, like so:

A, B, C ►

In these cases, one takes away the pointer symbol, replaces the commas with conjunctions (adding brackets to prevent ambiguity), puts brackets around the whole thing and adds a negation, like so:

$$\neg(A \wedge (B \wedge C))$$

If there is nothing on the left-hand-side but more than one thing on the right-hand-side, one simply replaces the commas with disjunctions, adding brackets to prevent ambiguity. So “► A, B, C” becomes:

$$(A \vee (B \vee C))$$

For sequents which are (a) of level 1 (b) not 'degenerate' (i.e. have no empty sides) and (c) free of set-names, the procedure is the following. One starts at the outer-scope pointer, and replaces that with a conditional. Then the commas separating the sentence names on the left-hand-side can be replaced with conjunction operators, adding brackets in any way that prevents ambiguity. The commas separating the sentence names on the right-hand-side can be replaced with disjunction operators, again adding brackets to prevent ambiguity. Take the following sentence:

$$P, Q, R \blacktriangleright S, T, U$$

This will become:

$$(P \wedge (Q \wedge R)) \rightarrow (S \vee (T \vee U))$$

If there are set names in the sequent, the way to translate it is to replace the set name with the names of its members, separated by commas, and then translating it in the above way.

These instructions cover the translation of any level 1 sequent. But there are also, of course, sequents of higher levels: sequents with embedded sequents, which in turn might have embedded sequents, etc. In those cases, one starts by translating the level 1 sub-sequents that appear in the sequent, and putting them between brackets. Having done that, one moves up a level. One translates those sub-sequents that are flanked by only level 0 and level 1 sequents. One does this by using the same procedures described above for level 1 sequents, treating the now-translated and bracketed level 1 sequents just as if they were level 0 sequents. One puts the results between brackets. By doing this, one can translate all the level 2 sub-sequents of the sequent. Then one moves up to the level 3 sequents, applying

the same procedure again, now treating all the level 1 and level 2 sequents just as if they were level 0 sequents. In this manner one goes up through the levels of embedding, for as long as it takes. Again, if the sequent has no outer pointer, then it will take forever to translate the sequent. But one will know how to do it.

And of course, if one wants to translate to a version of PL with a more limited vocabulary, say without the conjunction, one can simply translate the sequents in the manner just indicated, and then translate the obtained sentences to a more austere version of PL afterwards.

III.II.IV. Translating Sequentese: Recoverability

We've now seen how to translate PL into Sequentese and vice versa. But two-way translatability should not be all that we ask: we should also demand that a sentence of PL that gets translated into Sequentese and back again ends up being the *same* sequent as the sequent that we started with, or at least logically equivalent to it⁸². Our translation manuals are useless if they don't 'match up'. Can we guarantee that this happens? Let's take a look.

We start off by assuming a version of PL which has a negation, disjunction, a conditional and a conjunction. A sentence of PL is then either (1) an atomic sentence; (2) a negation; (3) a disjunction; (4) a conditional or (5) a conjunction. Let's look at these cases.

(1) The translation of an atomic sentence A of PL to Sequentese is a sequent with an empty left-hand-side ($A \blacktriangleright$). The translation of such a sequent into PL is the original atomic sentence P .

(2) A PL sentence the main connective of which is a negation ($\neg P$) gets translated as a sequent with an empty right-hand-side ($\blacktriangleright P$), no

82. Of course, we haven't yet specified what logical equivalence amounts to for sequents of Sequentese. We will do so below, but at the present juncture, it will not turn out to matter, for what we recover is the exact same sequent, and identity should suffice for logical equivalence, however logical equivalence is defined.

matter whether P is a molecular or atomic sentence. A sequent of that form gets translated back into PL as ' $\neg P$ '.

(3) A PL sentence the main connective of which is a disjunction (' $P \vee Q$ ') gets translated as a sequent with an empty left-hand-side and two things on the right-hand-side (' $\triangleright P, Q$ '). It does not matter whether P or Q are atomic or molecular. A sequent of that form gets translated back into PL as ' $P \vee Q$ '.

(4) A PL sentence the main connective of which is a conditional (' $P \rightarrow Q$ ') will get translated as a sequent with the antecedent on the left-hand-side and the consequent on the right-hand-side (' $P \triangleright Q$ ') no matter whether P and Q are atomic or molecular. Vice versa, a sequent with (proto-)sequents on both sides gets translated as a sentence of PL the main connective of which is a conditional.

(5) In the PL-to-Sequentese manual above, we didn't give a recipe for translating PL conjunctions, instead assuming a version of PL which didn't have any. Instead we suggested translating such sentences to the conjunction-free fragment of PL, and translating the result into Sequentese. The simplest way to translate a sentence of PL which has a conjunction as its main connective (' $P \wedge Q$ ') is to exploit De Morgan's equivalences: ' $P \wedge Q$ ' is equivalent to ' $\neg(\neg P \vee \neg Q)$ '. That would get translated into Sequentese as ' $(\triangleright(P \triangleright)), (Q \triangleright) \triangleright$ '. When we translate that back into PL using the Sequentese-to-PL manual, it becomes ' $\neg(\neg P \vee \neg Q)$ '. And this is, of course, De-Morgan-equivalent to ' $P \wedge Q$ '.

As one can tell from these considerations, the recoverability results are not sensitive to the potential complexity of the sub-sentences P and Q . If P and Q are atomic, then recoverability obtains for them. If P and Q are molecular sentences with a single level of complexity (their sub-sentences are atomic) then recoverability applies to them as well, given the above. And if, for a molecular sentence of any level of complexity, recoverability obtains for its immediate sub-sentences, then, given the above, recoverability will obtain for it. So, by

mathematical induction on the complexity of sentences, sentences of PL of any level of complexity, provided they 'bottom out', are recoverable. Which is the desired result.

IV. Reasoning in Sequentese

We've said quite a lot about Sequentese now, and how it relates to PL. Nevertheless all we've spoken about is single sequents of Sequentese (how they are formed, how they are interpreted) and how these relate to single sentences of PL. But there's more to the connectives than how they function in single sentences. Importantly, logical connectives play a role in determining how sentences of PL relate to other sentences of PL: whether they are jointly consistent or inconsistent, whether one implies the other, whether they are equivalent, etc. There are logical relations between sentences, and PL captures these as formal relations: the way two complex sentences of PL are built up out of simpler sentences, by means of the connectives, determines (many of) the logical relations that those two sentences will stand in to each other.

Sequentese can recapture these structures, even though it has no connectives. Complex propositional attitudes are structured, and their structures are reflected in the sequents that express them. We can identify formal relations between sequents that reflect certain relations between the attitudes they express. There are sets of complex attitudes such that, independently of the individual merits of the attitudes in question, one cannot rationally hold all of them, and some sets of complex attitudes are such that one cannot, again independently of the individual merits of the attitudes, rationally hold all of them without also holding some further attitude. These sorts of rational relations between attitudes, I propose, can be invoked to explain the logical relations that logically complex sentences stand in to each other. Or at least, as far as the argument in this chapter goes, they can be invoked to explain those logical relations that are captured in PL by the logical connectives and

their rules. Beyond those logical relations, there are also the logical relation captured by predicate logic – I'll put discussion of those off until chapter V.

Let's get into the frame of mind with an example. Suppose one had the attitudes expressed by the following two sequents:

(i) $P \triangleright Q$

(ii) $\triangleright P$

In other words, one (i) rules out accepting P whilst rejecting Q and (ii) rules out rejecting P. There are a number of ways of making one's attitudes to P and to Q fit with the attitude expressed by (i). One could accept P and accept Q; reject P and reject Q; and reject P and accept Q. But, given what (ii) says, one doesn't want to take either of the last two of these options. If one did, then one would be rejecting P while also refusing to reject P: that seems a clear failure of rationality. The only option left is to accept P and accept Q. Since that would involve accepting Q, it seems the attitudes above commit one to the attitude expressed by the following sequent:

(iii) $\triangleright Q$

For if one's only remaining option involves accepting Q, rejecting Q is not on the menu. It should be easy enough to see that, given the translation manual given in section III.II.III, the move from sequents (i) and (ii) to sequent (iii) represents an application of Modus Ponens. Adopting the attitudes expressed by (i) and (ii) constrains what further attitudes one ought to adopt, in the sense that if one were not to comply with these further attitudes (i.e. if one were to reject Q, in the example above) one would end up with an irrational combination of attitudes: one would be rejecting Q while simultaneously ruling out doing exactly that.

The notion of having to adopt certain complex attitudes in order to stave off irrationality is at the heart of the logical expressivist treatment of logical relations. By reflecting on the ways that attitudes with certain structures constrain each other, we can identify a number of 'inference rules' for Sequentese: moves from sets of sequents of certain forms to further sequents, such that the latter sequents express attitudes that one is rationally constrained to have if one has the former

attitudes. As it turns out, the inference rules that can be motivated this way will behave in classical ways. That is to say: if one translates the premises and conclusion of a valid argument in Sequentese into PL, one would obtain a valid argument in classically interpreted PL, and if one translated the premises and conclusion of a classically valid argument in PL into Sequentese, one would obtain a valid Sequentese argument.

This equivalence is, of course, exactly the result we want. In sections III.II.II – III.II.IV we’ve provided translation manuals for going back and forth between sentences of PL and sequents of Sequentese. But sequents of Sequentese do not derive their interpretation from their PL translations: we’ve provided an independent manual for rendering sequents into English descriptions of complex propositional attitudes, and these renderings are meant provide the interpretations of sequents. Similarly, the sentences of PL do not derive their interpretation from their translations into sequents of Sequentese. PL comes with its own standard classical model-theoretic semantics, which any logic textbook will provide one with. These model-theoretic semantics provide the rationale for the rules of the various systems of formal reasoning (natural deduction, truth tables, tableaux, sequent calculus) that have been devised for PL. Similarly, it is the interpretation of Sequentese in terms of complex propositional attitudes that provides the rationale for the sequent-calculus-style system of formal reasoning for Sequentese that I’ll present in sections IV.I – IV.III below. So if the Sequentese calculus turns out to capture a deductive consequence relation coextensive with that of PL, this is a substantive result: it’s not guaranteed by any link between the semantics of PL and the semantics of Sequentese, because the two semantics are independent, and the link between them hasn’t been established yet.

However, once it has been established that Sequentese and classical PL capture the same relation of deductive consequence, this is a result that’s philosophically useful. For that will allow us to claim that the interpretation of Sequentese, in terms of complex propositional attitudes – the logical expressivist semantics – provides a viable and non-revisionary semantics for classical PL. And that’s important because the logical expressivist semantics doesn’t have the metaphysical commitments that

the standard semantics of classical PL seems to have – there are no logically complex propositions in the interpretations of Sequentese.

The system of inference rules that I'll motivate for Sequentese will bear a general resemblance to the multiple-conclusion sequent calculus. The rules for the normal sequent calculus can be sorted into on the one hand a bunch of structural rules, not sensitive to the presence of any connectives in sequents, and a bunch of inference rules which are concerned with specific connectives – for every connective, an introduction and an elimination rule. The Sequentese calculus likewise has a bunch of structural rules – as it turns out, the same structural rules as the multiple-conclusion sequent calculus – but as Sequentese doesn't have any connectives, the further rules will be concerned with the behaviour of embedded sequents on either side of the pointer. There will be some 'embedding' and 'extraction' rules, which tell one how to move to and fro between sequents with and without certain forms of embedding.

IV.I. Motivating the Inference Rules

Below, I'll argue for a number of inference rules for the Sequentese formalism. Some of these will be completely general and correspond to standard structural rules of classical logic (and many non-classical logics besides): Transitivity, Reflexivity, Monotonicity (or Weakening) and Permutation. Then there will be six more inference rules which govern the embedding and un-embedding ('extraction', as I will call it) of sequents within other sequents. I will then proceed to argue that these rules get us a classical propositional calculus. I'll show that the inference rules of Sequentese, when we translate their premise- and conclusion-schemata into PL, correspond to classically inference-patterns. And I'll show that the axioms and inference rules of classical propositional logic (more specifically the axioms and inference rules of the Hilbert-style formalisation of classical propositional logic) are derivable in the Sequentese calculus.

But all that won't prove much about the viability of logical expressivism unless the inference rules which I lay down for Sequentese can be motivated in some way other than showing that they get us classical logic. They need to have motivations in distinctly expressivist terms. And to be in a position to provide motivations of that sort, I'll first present and argue for a bunch of general principles of synchronic rationality that I take to govern simple and complex propositional attitudes. These principles are not inference rules: they tell us that certain combinations of attitudes (describable in schematic terms) are irrational. These principles will then be used (in the next section) to argue for some principles of diachronic rationality, i.e. principles that govern transitions in thought. Those will be the inference rules of Sequentese.

All in all, six principles of synchronic rationality are in play, though they come in bunches and don't all need to be motivated individually. I'll present the principles, and then discuss what motivates them, and what issues arise with them. First off, I'll lay down two principles that govern acceptances and rejections of atomic propositions. For any proposition A:

BB: Don't accept A and reject A;

AA: Don't rule out accepting A and rule out rejecting A;

The first of these, which I call "BB" for "Bilateral Bi-Exclusion", is the most straightforward. It can be seen as the bilateralist propositional attitude analogue of the law of non-contradiction. In fact, this principle makes it the case that the law of non-contradiction comes out as a logical truth in Sequentese. That is not, however, to say that it entails the law of non-contradiction: some philosophers, such as Graham Priest, accept the above principle but reject the law of non-contradiction. (As noted in section I.I of this chapter, Priest uses it to give an account of disagreement that is consonant with dialetheism.) It is only in the particular context of logical expressivism that the above principle motivates the LNC.

The second principle, called "AA" for "Anti-Agnosticism", has a few more subtleties. Note that this principle is importantly different from the simpler principle 'either accept A, or reject A'. This latter principle would stand to the law of

excluded middle in the same way as BB stands to the law of non-contradiction. But that principle, while it would get us the results that we want, would be implausibly strong. There are clearly situations in which neither attitude is warranted: cases in which we don't have any evidence concerning A. What the principle above prohibits is what we might call come-what-may agnosticism: an attitude whereby one refuses to take any attitude to A, regardless of whether one has evidence bearing on it. What it does not prohibit is *pro tanto* agnosticism: not having an opinion on A because one has no reason to have one or the other attitude.

Why should come-what-may agnosticism be regarded as irrational? One might worry about the following. Suppose I know, of some proposition X, that X is unknowable: we are in principle unable to obtain any evidence that bears on the truth value of X. Shouldn't I, in this case, rule out accepting A and rejecting A? What's stopping me?

The issue here is that acceptance and rejection, as we explained these terms in section I.I.II., do not simply map onto belief and disbelief, but a bunch of cognate attitudes besides, not all of which are governed by the exact same norms. Belief and disbelief are attitudes that are evidence-constrained: one ought to believe what one has a certain amount of evidence for and not more than a certain amount against, and one ought to disbelieve what one has a certain amount of evidence against and not more than a certain amount against. If one knew that evidence for and against X was in principle unobtainable, then one could rule out believing X and rule out disbelieving X. But by ruling out accepting X and ruling out rejecting X, one rules out too much. One would rule out, for example, supposing that X and one would rule out counter-supposing that X. But whether or not evidence for or against X is obtainable, reasoning under the (counter-)supposition that X should still be fine. Likewise imagining that X or counter-imagining that X should also be fine. I cannot think of any feature that a proposition might have which render it in principle inappropriate for any form of acceptance and rejection, and render it appropriate to rule out accepting it and rule out rejecting it.

There are analogues of the principles above for complex attitudes whose objects are complex attitudes in their own right. For any pattern *P*:

ABB: Don't rule in and rule out P ;

AAA: Don't rule out ruling in P and rule out ruling out P ;

These principles are straightforward analogues of the ones above. The first one, called "ABB" for "Ascended Bilateral Bi-Exclusion", is again a bilateralist propositional attitude analogue of non-contradiction. The second one, called "AAA" for "Ascended Anti-Agnosticism" is again a principle prohibiting come-what-may agnosticism, this time concerning patterns rather than propositions.

Then there are principles which are, so to say, cross-level: they rule out certain combinations of attitudes and attitudes towards the former attitudes. For any pattern P :

H1: Don't rule out a pattern P while instantiating it.

H2: Don't rule in a pattern P without instantiating it, or instantiate it without ruling it in.

One might call these 'harmony principles' (hence the names). Basically, these rules tell us that our attitudes should conform to our attitudes towards those attitudes. If I am against some pattern of attitudes in the sense of ruling it out, it would be schizophrenic of me to have those very attitudes. Similarly, if I'm in favour of some pattern of attitudes, it would be schizophrenic of me not to have those attitudes. Note that there's a certain asymmetry between these principles: If I rule out a pattern, this just means I cannot rationally have *all* of the attitudes that make up that pattern. But if I rule a pattern in, I cannot rationally lack *any* of the attitudes in question. Another way to formulate that would be to say that ruling out a pattern is not *distributive*, in the sense that it means that I thereby rule out every individual attitude in the pattern, but ruling in is distributive. If I rule a pattern in, I thereby rule in all of the individual attitudes making up the pattern.

How are all of the above principles to be justified? The non-contradiction analogues, I believe, don't need much in the way of justification. Violations of these principles seem to be to be paradigmatic cases of failures of rationality. If it is not irrational to violate BB and ABB, then I don't know what rationality is. Note that while this principle does help motivate the structural rules and inference rules of

Sequentese in such a way that the resulting logic is classical, the principle itself is one that's uncontroversial among bilateralists, including those who advocate paraconsistent logics.

Justification-wise, the harmony principles seem almost as straightforward to me as the bi-exclusion principles. If one has an attitude but also refuses to have it, or if one lacks an attitude but also endorses having that same attitude, something is clearly off. Some schizophrenia obtains, and one or the other attitude should go.

The anti-agnosticism principles are less straightforwardly justified. They seem to be the ones that are most logically contentious: advocates of paracomplete logics will presumably not accept them, because they think that there are propositions which are neither true nor false, and thus are neither appropriate candidates for acceptance, nor for rejection. It should be remembered, of course, that I'm not in the business, in this chapter, of defending classical logic: I'm in the business of showing that a classical logical expressivism is there to be had. In order to get it, we need principles AA and AAA. But as long as anyone willing to endorse classical logic is also willing to endorse AAA, that's fine.

The disagreement between classical logic and paracomplete logics is, on some level, a metaphysical one, and more specifically a disagreement about the metaphysics of truth and falsity. Classicists generally maintain that these two properties partition the realm of propositions exclusively and exhaustively; the paracomplete logician denies that they do so exhaustively. These metaphysical positions may then be further explained in terms of what truth and falsehood are. An intuitionist, for instance, will hold that they are to be understood ultimately in epistemic terms, and on the basis of this will claim that truth and falsehood do not outrun knowability: any truth can be known to be true, and any falsehood can be known to be false. Since for some propositions we cannot know their truth-value (or so the intuitionist holds) some propositions are neither true nor false.

But to get AA and AAA, one need not enter into this debate. The expressivist needn't claim that truth and falsity exhaustively partition the realm of proposition. For, as noted, the principles above govern acceptance and rejection, not just belief and disbelief and certainly not just knowledge-of-truth and knowledge-of-

falsehood. They govern a family of attitudes of taking-a-stance-on, which include attitudes like supposing-that, which is not governed by evidence (or, indeed, truth) in the way that belief arguably is. If there are propositions that are truth-value-less, or merely unknowable, then one isn't under any obligation to accept them or reject them. After all, one needn't believe or disbelieve them, in the absence of evidence or truth-value, and one needn't have any suppositions or imaginings concerning them either (since such attitudes are never compulsory anyway). And AA and AAA do not imply that one has such obligations. AA and AAA prohibit, however, the joint ruling-out of acceptance and rejection, for any proposition, and the joint ruling-out of ruling-in and ruling-out, for any pattern. And that holds for the truth-value-less and unknowable propositions as well, if there are such things. After all, whatever the epistemic status of a proposition, one ought to allow oneself to imagine it to be the case, or imagine it not to be the case, or suppose it to be the case, or suppose it not to be the case. By ruling all of that out, one rules out too much.

Similar considerations would apply in the case of *indeterminate* propositions: though perhaps one ought not accept or reject an indeterminate proposition, one is still at liberty to take one of these other types of stances on such a proposition, and to rule that out would be illegitimate.

One might object, at this point, that AA and AAA only hold for the expressivist because they help themselves to these unusually wide notions of acceptance and rejection, covering the whole family of stance-taking propositional attitudes. If one did not make this somewhat unusual choice, the desired results wouldn't obtain. The choice of these inclusive notions of acceptance and rejection is not, however, merely *ad hoc*. Logic is meant to provide us with principles that govern our reasoning, not just principles which govern our beliefs and disbeliefs. One can reason about things which one merely supposes or merely imagines, and the rules of reasoning aren't any different in those cases. A 'logic' which only governs which beliefs and disbeliefs I may combine with other beliefs and disbeliefs does not do all of what we conventionally think a logic ought to do. So we *should* be taking into account these other sorts of attitudes from the get-go, given that they are ultimately to be governed by the same logical principles.

One might be getting a little curious at this point as to what would happen if we *were* to lose principles AA and AAA. Since these principles don't look attractive to a paracomplete logician, might that mean that if we drop them, we get a recognisable paracomplete logic out of it? As a matter of fact, we do get a paracomplete logic, but not a familiar one. The reason is that (as we'll see below) AA and AAA are involved in motivating the structural rule of *Transitivity*. If we drop AA and AAA, the resultant logic would have a non-transitive consequence relation. Non-transitive logics are very weak logics, and it is hard to believe that such a logic could be the one that governs our thinking.⁸³ That also means, of course, that if we ever try to adapt the logical expressivist proposal to paracomplete logics, the adaptation won't be very straightforward, and a major challenge will be finding some way of motivating *Transitivity* which doesn't rely on prior principles which are anathema to paracomplete logicians. But that is a problem for another day.

These, then, are the principles of synchronic rationality which will be involved in supplying Sequentese with some inference rules.

IV.II. Motivating the Structural Rules

On the basis of the rationality principles laid down in the previous section, we can motivate a few general claims about how sequents relate to other sequents⁸⁴, regardless of whether these sequents display any embedding of sequents. This will amount to showing that the pointer of Sequentese has the same structural features as the PL consequence relation: it obeys permutation and simplification, it's monotonic, transitive and reflexive.

It might be worth saying something about the general procedure I'll be employing. The principles of synchronic rationality above are impermissions: they tell us not to do certain things. Inference rules are, strictly speaking, permissions:

83. But for a dissenting voice, see Ripley (forthcoming(a) & forthcoming(b))

84. In the following I'll be relying again on some work by Greg Restall: his (2005), (2008) and (forthcoming). My way of motivating the rules is a bit different from his, but the general idea of motivating these rules in terms of rational constraints on acceptance and rejection I take from his work.

they allow us to make certain transitions in thought. How might the former relate to the latter? Well, when one sequent S2 *follows from* or *is derivable from* some other sequent S1, that means that if one instantiates (in some way or other) S1, also instantiating S2 won't make one (synchronically) irrational. Showing that S2 follows from S1 thus involves showing that adding the attitude expressed by S2 to a stock of attitudes that includes S1 doesn't lead to any problems. But it shows a little more than that too. Attitudes of acceptance, rejection, ruling-in and ruling-out come in 'normative pairs', as it turns out. When some attitudes license a further attitude, there will be an attitude opposed to that further attitude that will not be rationally compatible with the former attitudes. More concretely, if the acceptance of P is licensed, then the rejection of P is not, and vice versa, and if ruling-in some pattern is licensed, then ruling it out is not, and vice versa. So along with the permissions also come impermissions.

On to the rules, then. Below, capital Roman letters can symbolise both atomic propositions and embedded sequents. In many cases, in explaining the rules, I will talk just of 'the acceptance of A' or 'the rejection of B' when illustrating the rules. It should be understood that these letters can be taken to stand for sequents, and in those cases, the attitudes talked about are attitudes of ruling in and ruling out patterns, rather than attitudes of acceptance and rejection. But the explanations would quickly become unreadable if this were made explicit at every point.

Permutation. Note first that patterns, the objects of rulings-in and rulings-out, do not have any ordering to them. They are simply sets of acceptances, rejections, rulings in and rulings out, and not ordered sets. When we write down one of these patterns, we must of course write down the members of it in some order or other; but that order is irrelevant to the individuation of the pattern. If we had written down those members in a different order, we'd have described the same patterns. So, to illustrate, if the pattern of accepting A, accepting B and rejecting C is refused, this is the same as refusing the pattern of accepting A, rejecting C and accepting B. (Of course, what I *cannot* change without changing the identity of pattern is which propositions are marked for acceptance and which are marked for rejection. Those features do matter for the individuation of patterns and complex attitudes.) Let's

see what the upshot of this is for the Sequentese calculus. The attitude expressed by:

$$A, B \triangleright C$$

Is the same as the attitude expressed by:

$$B, A \triangleright C$$

Here, by switching around the A and the B, we've changed nothing about the attitude expressed. So here's one formal relation between sequents of Sequentese: for any sequent S , a sequent S^* that results from either changing the order of the arguments on S 's left-hand-side follows from S , and vice versa. Of course, since sequents S and S^* say exactly the same thing, this doesn't amount to anything more than saying, tautologically, that one is rationally permitted to have the attitudes that one is rationally permitted to have. This also means that if one were to have the attitude expressed by S but ruled in the pattern that S^* rules out, one would be in a state of irrationality: one would be ruling out and ruling in the same pattern, in violation of principle ABB.

The same applies, of course, when we're dealing with a premise-sequent with multiple propositions on the *right-hand-side* of the pointer, i.e. a sequent which expresses the refusal of a pattern involving multiple rejections. The order of these rejections is of no import, as long as they stay on the right side of the pointer. So for any sequent S , a sequent S^* that results from either changing the order of the arguments on S 's right-hand-side is a consequence of S , and vice versa. So, to present both of these derivability claims formally:

$$\Gamma, A, B, \Delta \triangleright C$$

$$\Gamma, B, A, \Delta \triangleright C$$

$$A \triangleright \Gamma, B, C, \Delta$$

$$A \triangleright \Gamma, C, B, \Delta$$

and

These correspond to a standard structural rule of classical (multiple-conclusion) sequent calculus: *Permutation*.

Weakening. If a bunch of attitudes are part of a pattern that one rules out, and one does in fact have all of the attitudes in the pattern, one is in an irrational state: one is in violation of one or more of the harmony rules. One might alleviate this irrationality by dropping one of the attitudes in question. But importantly, one could never do so by *adding* any attitudes to one's stock. So if one rules out a pattern P, one cannot, on pain of irrationality, rule in any superset of P (call it a *super-pattern* of P). For to do so would be to, in effect, rule in P, and that would be a violation of ABB. One can, however, rule out the super-pattern.⁸⁵

To illustrate: if one ruled out accepting A and rejecting B, one cannot rationally rule in accepting A, rejecting B and rejecting some further proposition C. By doing so one would *ipso facto* rule in accepting A and rejecting B, and that would violate ABB. But one could rule out this super-pattern, and in fact that's the only attitude one would be allowed to have to that pattern. So, summing up, as regards the derivability relation of Sequentese we can say that a sequent S implies any sequent S* which results from S by adding the name of a proposition, or the name of a set of propositions, or a sequent to either side of that sequent. Formally:

$$\begin{array}{ccc}
 \Gamma \blacktriangleright \Delta & & \Gamma \blacktriangleright \Delta \\
 \text{-----} & \text{and} & \text{-----} \\
 \Gamma, A \blacktriangleright \Delta & & \Gamma \blacktriangleright A, \Delta
 \end{array}$$

This corresponds to *Weakening*, another standard structural rule of classical sequent calculus. Or, equivalently, we can say that the relation expressed by the pointer of Sequentese is *monotonic*.

Transitivity. Suppose that one rules out accepting A and rejecting B (i.e. what would be expressed by $A \blacktriangleright B$) and suppose that one also rules out accepting B and rejecting C (i.e. $B \blacktriangleright C$). If one does accept A, one ought not then reject B, on pain of violating H1. If one is going to form an attitude to B, it ought to be accepting B. If

85. Cf. Restall (ibid.).

one does accept B, one ought not then reject C, again on pain of violating H1. If one is going to have any attitude toward C, it can only be acceptance.

The acceptance of A isn't directly incompatible with the rejection of C. But if one accepts A this will commit one to not rejecting B. In other words, to combine ruling out accepting A and rejecting B with actually accepting A is to in effect rule out the rejection of B. Similarly, if one rejects C, this will commit one to not accepting B. In other words, to combine ruling out accepting B and rejecting C with actually rejecting C is to in effect rule out the acceptance of B. The upshot is that one would rule out the acceptance of B and the rejection of B, in violation of principle AA. So one ought not both accept A and reject C. Formally:

A ► B

B ► C

A ► C

And the same would apply if we filled in sequents for A, B and C, albeit that the synchronic rationality principle invoked would be AAA. The above rule corresponds to *Transitivity*, a third standard structural rule of classical sequent calculus.

Reflexivity. BB rules out accepting a proposition A and rejecting it, and ABB rules out ruling in a pattern P and ruling it out. This means, reading "A" as a schematic letter ranging both over propositions and sequents:

A ► A

This corresponds to the standard structural rule called *Reflexivity*.

Contraction. Patterns are sets of attitudes to propositions. Sets have the property that they can only contain any given item once. That means that there is no pattern of accepting A and accepting A, distinct from the pattern of accepting A. There is also no pattern of rejecting B and rejecting B, distinct from the pattern of rejecting B. So the following moves would take one from a sequent expressing an attitude to a sequent expressing that same attitude more succinctly:

$$\begin{array}{ccc}
 \Gamma, A, A \blacktriangleright B, \Delta & & \Gamma, A \blacktriangleright B, B, \Delta \\
 \text{-----} & \text{and} & \text{-----} \\
 \Gamma, A \blacktriangleright B, \Delta & & \Gamma, A \blacktriangleright B, \Delta
 \end{array}$$

This corresponds to a structural rule sometimes referred to as *Contraction*, or *Simplification*.

It seems, then, that just from the nature of the attitudes expressed and the synchronic rationality claims that one can plausibly make about them, one can motivate rules for reasoning with ‘ \blacktriangleright ’ which make it behave just as we would expect the turnstile of multiple-conclusion sequent calculus to behave.

These rules, however, only describe some of the logical behaviour of the pointer. They don’t allow us to do anything interesting with embedded sequents. We need some more rules for that, rules which are not analogues to familiar structural rules.

IV.III. Embedding and Extraction Rules for Sequentese

There are two ways to embed a sequent: one can embed it on the left-hand-side, and one can embed it on the right-hand-side. Here is a proposal for a rule governing left-hand-side embedding:

$$\begin{array}{l}
 \text{(i) } \Gamma \blacktriangleright A, \Delta \\
 \text{(ii) } \Gamma, B \blacktriangleright \Delta \\
 \text{-----} \quad \text{LEM}
 \end{array}$$

(iii) $\Gamma, (A \blacktriangleright B) \blacktriangleright \Delta$

What's the thinking here? Well, let's suppose Γ and Δ name sets of attitudes to propositions (the former a set of acceptances, the latter a set of rejections). The status of these particular attitudes, let's stipulate, is not presently up for debate. They're part of the background, so to say. Taking these as fixed, (i) expresses that one rules out rejecting A. And likewise, taking these as fixed, (ii) expresses that one rules out accepting B.

Principle H1 prohibits having the attitude expressed by (i) and rejecting A. It also prohibits having the attitude expressed by (ii) and accepting B. So the only rationally permissible way to extend one's attitudes towards A and B would be to accept A, and reject B. But suppose one were to rule out accepting A and rejecting B (i.e. have the attitude expressed by ' $A \blacktriangleright B$ '). Then one would be in violation of H1. So ruling out accepting A and rejecting B is not an option, given (i) and (ii). So the attitude one ought to have to that attitude is that of ruling it out. This is what one expresses by embedding the sequent ' $A \blacktriangleright B$ ' on the left-hand-side.

That's a rule which leads one from sequents without embedding to a sequent with embedding. So let's call that an *embedding* rule, more precisely the left-hand-side embedding rule, or LEM.

If that's the left-hand-side embedding rule, what would the right-hand-side rule(s) be? I propose:

(i) $\Gamma, A \blacktriangleright \Delta$

----- REM(r)

(ii) $\Gamma \blacktriangleright (A \blacktriangleright B), \Delta$

(i) $\Gamma \blacktriangleright B, \Delta$

----- REM(a)

(ii) $\Gamma \blacktriangleright (A \blacktriangleright B), \Delta$

These can be motivated fairly straightforwardly, as follows. Suppose I rule out accepting A (given Γ and Δ), as expressed by the leftmost sequent (i). Then *a fortiori* I rule out doing this while rejecting B, which is what the leftmost sequent

(ii) expresses. Similarly, suppose I rule out rejecting B (given Γ and Δ) as the rightmost sequent (i) expresses. Then *a fortiori* I rule out doing this while accepting A, which is what the rightmost sequent (ii) expresses.

Those are the three embedding rules that we need. Call a rule which takes one from a sequent with embedding to a sequent with no embedding (or less embedding) an *extraction* rule. Of these there will be left- and right-hand-side versions as well. What would be the left-hand-side extraction rule(s)? I propose:

$$\begin{array}{ccc}
 \text{(i) } \Gamma, (A \blacktriangleright B) \blacktriangleright \Delta & & \text{(i) } \Gamma, (A \blacktriangleright B) \blacktriangleright \Delta \\
 \text{-----} & \text{LEX(r)} & \text{-----} & \text{LEX(a)} \\
 \text{(ii) } \Gamma, B \blacktriangleright \Delta & & \text{(ii) } \Gamma, \blacktriangleright A, \Delta
 \end{array}$$

These rules can be most easily motivated by appealing to REM(r) and REM(a), above. Let's say that one refuses to refuse to accept A and reject B, as the premises marked (i) have it. There are, in the abstract, two ways of extending one's opinions to A and to B. Suppose, for reductio, that it were fine to also accept B. Then, by REM(a), above, one could infer a sequent which had the sequent "A \blacktriangleright B" embedded on the right-hand-side. But if one did that, one would be in violation of ABB, by ruling out ruling out "A \blacktriangleright B", and ruling out ruling in "A \blacktriangleright B". So it's not in fact fine to also accept B. Accepting B ought to be ruled out, and so "B" goes on the left-hand-side, as line (ii) of the LEX(r) rule has it. Similarly, if it were okay to reject A, then by REM(r) it should be okay to infer a sequent that has "A \blacktriangleright B" embedded on the left-hand-side. But then one would again be in violation of ABB. So it's not okay to reject A, and rejecting A ought to be ruled out, as (ii) of the LEX(a) rule has it.

As a right-hand-side extraction rule, I propose:

$$\begin{array}{ccc}
 \text{(i) } \Gamma \blacktriangleright (A \blacktriangleright B), \Delta & & \\
 \text{-----} & \text{REX} &
 \end{array}$$

(ii) $\Gamma, A \blacktriangleright B, \Delta$

The motivation is very basic in this case. If one rules out ruling in the pattern of accepting A and rejecting B, as sequent (i) expresses, then one ought not accept A and reject B. For, if one did, one would have to rule in accepting A and rejecting B, by principle H2. And then one would be instantiating a pattern (ruling in accepting A and rejecting B) one rules out, in violation of H1. So one ought to rule out accepting A and rejecting B, as sequent (ii) expresses.

Those are the three extraction rules of Sequentese. Those three, together with the three embedding rules and the five structural rules set out earlier, are all the rules one needs to derive everything that one can derive the translation of in PL. Also, we can show that these rules don't make Sequentese any stronger than PL: all of these rules are such that they translate to valid arguments of PL.

IV.IV. Is Sequentese Classical?

To show that Sequentese is a classical language, we'll have to show that it is a) as powerful as classical logic (i.e. that everything that can be derived in a classical system can be derived in Sequentese) and that it is no more powerful (i.e. that everything that can be derived in Sequentese can also be derived in classical logic). To make the task a little clearer, let's introduce a new symbol: " \gg ", the *metapointer*. Let " $\gg S$ " mean that there is a proof in Sequentese of the sequent S from no premises, and let " $S1 \gg S2$ " mean that there is a proof of the sequent S2 in Sequentese from the sequent S1 and no others, and so on. The metapointer " \gg " can have any number of sequents on its left-hand-side, but unlike the pointer, it can only have a single sequent on its right-hand-side.

What exactly is a proof of Sequentese? Let's stipulate that it is a series of well-formed sequents of Sequentese such that each sequent is either (i) a premise or (ii) obtained from zero (in the case of Reflexivity) or one or more of these premises by

one of the inference rules we've introduced in the previous two sections. The last sequent of such a series is the conclusion of the proof: in the notation above, it's the sole sequent on the right-hand-side of the metapointer.

Furthermore, let " $T_{\text{SEQ}}(\phi)$ " mean "the translation of ϕ into Sequentese", where ϕ is a well-formed formula of PL. What we want to show, then, is the following (let the turnstile indicate classical derivability):

$$\phi \vdash \psi \text{ iff } T_{\text{SEQ}}(\phi) \gg T_{\text{SEQ}}(\psi)$$

This of course requires us to show that:

$$\text{If } \phi \vdash \psi, \text{ then } T_{\text{SEQ}}(\phi) \gg T_{\text{SEQ}}(\psi)$$

And:

$$\text{If } T_{\text{SEQ}}(\phi) \gg T_{\text{SEQ}}(\psi), \text{ then } \phi \vdash \psi.$$

I'll argue for the topmost conditional first, in the next section, and then for the bottom conditional in the section after that.

IV.IV.I. Classical Derivability Implies Sequentese Derivability

I will take the following approach to showing that $\phi \vdash \psi \rightarrow T_{\text{SEQ}}(\phi) \gg T_{\text{SEQ}}(\psi)$: using the Hilbert-style axiomatisation of classical propositional logic, I'll show that the Sequentese translations of each of the three axiom schemata of classical propositional are derivable from no premises using the inference rules of Sequentese, and that its one inference rule (modus ponens) also remains valid in its Sequentese translation.⁸⁶ Any proof in PL can be done by using only these axioms and one inference rule, and so one could reconstruct a (lengthy) Sequentese version of any such PL proofs by inserting one of the proofs given below where the PL proof appeals to one of these axioms or to the inference rule. The three axiom schemata are:

86. I'm using Sider (2010) as my source for Hilbert-style classical logic.

PL1: $\phi \rightarrow (\psi \rightarrow \phi)$

PL2: $(\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi))$

PL3: $(\neg\psi \rightarrow \neg\phi) \rightarrow ((\neg\psi \rightarrow \phi) \rightarrow \psi)$.

And the inference rule is Modus Ponens:

(i) $\phi \rightarrow \psi$

(ii) ϕ

----- MP

(iii) ψ .

Let's proceed to the proofs, then.

PL1: $\phi \rightarrow (\psi \rightarrow \phi)$

$T_{SEQ}((\phi \rightarrow (\psi \rightarrow \phi))) = \phi \blacktriangleright (\psi \blacktriangleright \phi)$

We need to show that: $\gg \phi \blacktriangleright (\psi \blacktriangleright \phi)$

Proof:

(1) $\phi \blacktriangleright \phi$

Reflexivity

(2) $\phi \blacktriangleright (\psi \blacktriangleright \phi)$

REM(a): 1

QED

PL2: $(\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi))$

$T_{SEQ}((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi))) = (\phi \blacktriangleright (\psi \blacktriangleright \chi)) \blacktriangleright ((\phi \blacktriangleright \psi) \blacktriangleright (\phi \blacktriangleright \chi))$

We need to show that: $\gg (\phi \blacktriangleright (\psi \blacktriangleright \chi)) \blacktriangleright ((\phi \blacktriangleright \psi) \blacktriangleright (\phi \blacktriangleright \chi))$

Proof:

(1) $(\phi \blacktriangleright (\psi \blacktriangleright \chi)) \blacktriangleright (\phi \blacktriangleright (\psi \blacktriangleright \chi))$

Reflexivity

(2) $(\phi \blacktriangleright (\psi \blacktriangleright \chi)), \phi \blacktriangleright (\psi \blacktriangleright \chi)$

REX: 1

- (3) $(\phi \triangleright (\psi \triangleright \chi)), \phi, \psi \triangleright \chi$ REX: 2
- (4) $(\phi \triangleright (\psi \triangleright \chi)), \phi, \psi \triangleright (\phi \triangleright \chi)$ REM(a): 3
- (5) $(\phi \triangleright (\psi \triangleright \chi)), \psi \triangleright (\phi \triangleright \chi), (\phi \triangleright \chi)$ REM(r): 4
- (6) $(\phi \triangleright (\psi \triangleright \chi)), \psi \triangleright (\phi \triangleright \chi)$ Simplification: 5
- (7) $(\phi \triangleright (\psi \triangleright \chi)), \psi \triangleright ((\phi \triangleright \chi) \triangleright (\phi \triangleright \chi))$ REM(a): 6
- (8) $((\phi \triangleright \chi) \triangleright (\phi \triangleright \chi)) \triangleright ((\phi \triangleright \chi) \triangleright (\phi \triangleright \chi))$ Reflexivity
- (9) $(\phi \triangleright \chi) \triangleright ((\phi \triangleright \chi) \triangleright (\phi \triangleright \chi))$ LEX(r): 8
- (10) $\triangleright \phi, ((\phi \triangleright \chi) \triangleright (\phi \triangleright \chi))$ LEX(a): 9
- (11) $(\phi \triangleright (\psi \triangleright \chi)), (\phi \triangleright \psi) \triangleright ((\phi \triangleright \psi) \triangleright (\phi \triangleright \chi)), ((\phi \triangleright \psi) \triangleright (\phi \triangleright \chi))$ LEM: 7, 10
- (12) $(\phi \triangleright (\psi \triangleright \chi)), (\phi \triangleright \psi) \triangleright ((\phi \triangleright \psi) \triangleright (\phi \triangleright \chi))$ Simplification: 11
- (13) $(\phi \triangleright (\psi \triangleright \chi)) \triangleright ((\phi \triangleright \psi) \triangleright (\phi \triangleright \chi)), ((\phi \triangleright \psi) \triangleright (\phi \triangleright \chi))$ REM(r): 12
- (14) $(\phi \triangleright (\psi \triangleright \chi)) \triangleright ((\phi \triangleright \psi) \triangleright (\phi \triangleright \chi))$ Simplification: 13

QED

PL3: $(\neg\psi \rightarrow \neg\phi) \rightarrow ((\neg\psi \rightarrow \phi) \rightarrow \psi)$

$T_{SEQ}((\neg\psi \rightarrow \neg\phi) \rightarrow ((\neg\psi \rightarrow \phi) \rightarrow \psi)) = ((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright (((\psi \triangleright) \triangleright \phi) \triangleright \psi)$

We need to show that: $\triangleright \triangleright ((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright (((\psi \triangleright) \triangleright \phi) \triangleright \psi)$

Proof:

- (1) $((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright ((\psi \triangleright) \triangleright (\phi \triangleright))$ Reflexivity
- (2) $((\psi \triangleright) \triangleright (\phi \triangleright)), (\psi \triangleright) \triangleright (\phi \triangleright)$ REX: 1
- (3) $((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright (\phi \triangleright), \psi$ LEX(a): 2
- (4) $((\psi \triangleright) \triangleright (\phi \triangleright)), \phi \triangleright \psi$ REX: 3
- (5) $((\psi \triangleright) \triangleright (\phi \triangleright)), ((\psi \triangleright) \triangleright \phi) \triangleright \psi$ LEM: 4
- (6) $((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright (((\psi \triangleright) \triangleright \phi) \triangleright \psi), \psi$ REM(r): 5

(7) $((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright (((\psi \triangleright) \triangleright \phi) \triangleright \psi), (((\psi \triangleright) \triangleright \phi) \triangleright \psi)$ REM(a): 6

(8) $((\psi \triangleright) \triangleright (\phi \triangleright)) \triangleright (((\psi \triangleright) \triangleright \phi) \triangleright \psi)$ Simplification: 7

QED

Modus Ponens: $\phi \rightarrow \psi, \phi \vdash \psi$

$T_{SEQ}(\phi \rightarrow \psi) = \phi \triangleright \psi$

We need to show that: $\phi \triangleright \psi, \triangleright \phi \gg \psi$

Proof:

- | | | |
|-----|----------------------------|--------------------|
| (1) | $\phi \triangleright \psi$ | Premise |
| (2) | $\triangleright \phi$ | Premise |
| (3) | $\triangleright \psi$ | Transitivity: 1, 2 |

QED

IV.IV.II. Sequentese Derivability Implies Classical Derivability

For the second task, that of showing that $\gg T_{SEQ}(\phi) \rightarrow \vdash \phi$, I will limit myself to showing that those rules that are specific to Sequentese (i.e. the embedding and extraction rules) are valid under translation. Sequentese also has another set of rules (Permutation, Weakening, Contraction, Reflexivity and Transitivity) that are analogues of structural rules of classical logic. The proofs of those are sufficiently trivial that I omit them here. So the rules I will show are valid are:

LEM: $(\Gamma \triangleright A, \Delta), (\Gamma, B \triangleright \Delta) \gg \Gamma, (A \triangleright B) \triangleright \Delta$

LEX(r): $(\Gamma, (A \triangleright B) \triangleright \Delta) \gg \Gamma, B \triangleright \Delta$

LEX(a): $\Gamma, (A \triangleright B) \triangleright \Delta \gg \Gamma, \triangleright A, \Delta$

REM(r): $\Gamma, A \triangleright \Delta \gg \Gamma \triangleright (A \triangleright B), \Delta$

REM(a): $\Gamma \triangleright B, \Delta \gg \Gamma \triangleright (A \triangleright B), \Delta$

REX: $\Gamma \triangleright (A \triangleright B), \Delta \gg \Gamma, A \triangleright B, \Delta$

I'll use " $T_{PL}(S)$ " to mean "the translation of sequent S into PL". By " $\wedge[\Gamma]$ " I mean "the conjunction of the members of Γ " and by " $\vee[\Gamma]$ " I mean the disjunction of the

members of Γ . For the proofs I'll use semantic tableaux or natural-deduction proofs, whatever proves the most convenient.

LEM: $(\Gamma \triangleright A, \Delta), (\Gamma, B \triangleright \Delta) \gg (\Gamma, (A \triangleright B) \triangleright \Delta)$

$T_{PL}(\Gamma \triangleright A, \Delta) = \wedge[\Gamma] \rightarrow (A \vee \vee[\Delta])$

$T_{PL}(\Gamma, B \triangleright \Delta) = (\wedge[\Gamma] \wedge B) \rightarrow \vee[\Delta]$

$T_{PL}(\Gamma, (A \triangleright B) \triangleright \Delta) = (\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$

So we need to show that $\wedge[\Gamma] \rightarrow (A \vee \vee[\Delta]), (\wedge[\Gamma] \wedge B) \rightarrow \vee[\Delta] \vdash (\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$

Proof (by Natural Deduction):

(1) $\wedge[\Gamma] \rightarrow (A \vee \vee[\Delta])$	Premise
(2) $(\wedge[\Gamma] \wedge B) \rightarrow \vee[\Delta]$	Premise
(3) $\wedge[\Gamma] \wedge (A \rightarrow B)$	Assumption
(4) $\wedge[\Gamma]$	3 \wedge -E
(5) $A \vee \vee[\Delta]$	1,4 \rightarrow -E
(6) $A \rightarrow B$	3 \wedge -E
(7) A	Assumption
(8) B	6,7 \rightarrow -E
(9) $\wedge[\Gamma] \wedge B$	4, 8 \wedge -I
(10) $\vee[\Delta]$	2,9 \rightarrow -E
(11) $\vee[\Delta]$	Assumption
(12) $\vee[\Delta]$	5, 7, 10, 11, 11 \vee -E
(13) $(\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$	3, 12 \rightarrow -I

QED

LEX(r): $(\Gamma, (A \triangleright B) \triangleright \Delta) \gg \Gamma, B \triangleright \Delta$

$T_{PL}(\Gamma, (A \triangleright B) \triangleright \Delta) = (\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$

$T_{PL}(\Gamma, B \triangleright \Delta) = (\wedge[\Gamma] \wedge B) \rightarrow \vee[\Delta]$

So we need to show that $(\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta] \vdash (\wedge[\Gamma] \wedge B) \rightarrow \vee[\Delta]$

Proof (by Natural Deduction):

(1) $(\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$	Premise
(2) $\wedge[\Gamma] \wedge B$	Assumption
(3) $\wedge[\Gamma]$	2 \wedge -E
(4) B	2 \wedge -E
(5) A	Assumption
(6) $A \rightarrow B$	4, 5 \rightarrow -I
(7) $\wedge[\Gamma] \wedge (A \rightarrow B)$	3, 6 \wedge -I
(8) $\vee[\Delta]$	1, 7 \rightarrow -E
(9) $(\wedge[\Gamma] \wedge B) \rightarrow \vee[\Delta]$	2, 8 \rightarrow -I

QED

LEX(a): $\Gamma, (A \blacktriangleright B) \blacktriangleright \Delta \gg \Gamma, \blacktriangleright A, \Delta$

$T_{PL}(\Gamma, (A \blacktriangleright B) \blacktriangleright \Delta) = (\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$

$T_{PL}(\Gamma, \blacktriangleright A, \Delta) = \wedge[\Gamma] \rightarrow (A \vee \vee[\Delta])$

So we need to show that $(\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta] \vdash \wedge[\Gamma] \rightarrow (A \vee \vee[\Delta])$

Proof (by semantic tableau):

(1) $(\wedge[\Gamma] \wedge (A \rightarrow B)) \rightarrow \vee[\Delta]$	Premise
(2) $\neg[\wedge[\Gamma] \rightarrow (A \vee \vee[\Delta])]$	\neg Conclusion
(3) $\wedge[\Gamma]$	$\neg \rightarrow$: 2
(4) $\neg[(A \vee \vee[\Delta])]$	$\neg \rightarrow$: 2
(5) $\neg A$	$\neg \vee$: 4
(6) $\neg \vee[\Delta]$	$\neg \vee$: 4
(7) $\neg[\wedge[\Gamma] \wedge (A \rightarrow B)]$ $\vee[\Delta]$ (contradiction!)	\rightarrow : 1
(8) $\neg \wedge[\Gamma]$ (contradiction!) $\neg(A \rightarrow B)$	$\neg \wedge$: 7

(9) A (contradiction!) $\neg\rightarrow$: 8

QED

REM(r): $\Gamma, A \blacktriangleright \Delta \gg \Gamma \blacktriangleright (A \blacktriangleright B), \Delta$

$T_{PL}(\Gamma, A \blacktriangleright \Delta) = (\wedge[\Gamma] \wedge A) \rightarrow \vee[\Delta]$

$T_{PL}(\Gamma \blacktriangleright (A \blacktriangleright B), \Delta) = \wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$

So we need to show that $(\wedge[\Gamma] \wedge A) \rightarrow \vee[\Delta] \vdash \wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$

Proof (by semantic tableau):

- | | | |
|------|--|---|
| (1) | $(\wedge[\Gamma] \wedge A) \rightarrow \vee[\Delta]$ | Premise |
| (2) | $\neg[\wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])]$ | \neg Conclusion |
| (3) | $\wedge[\Gamma]$ | $\neg\rightarrow$: 2 |
| (4) | $\neg((A \rightarrow B) \vee \vee[\Delta])$ | $\neg\rightarrow$: 2 |
| (5) | $\neg(A \rightarrow B)$ | $\neg\vee$: 4 |
| (6) | $\neg\vee[\Delta]$ | $\neg\vee$: 4 |
| (7) | A | $\neg\rightarrow$: 5 |
| (8) | $\neg B$ | $\neg\rightarrow$: 6 |
| (9) | $\neg(\wedge[\Gamma] \wedge A)$ | $\vee[\Delta]$ (contradiction!) \rightarrow : 1 |
| (10) | $\neg\wedge[\Gamma]$ (contradiction!) | $\neg A$ (contradiction!) $\neg\wedge$: 9 |

QED

REM(a): $\Gamma \blacktriangleright B, \Delta \gg \Gamma \blacktriangleright (A \blacktriangleright B), \Delta$

$T_{PL}(\Gamma \blacktriangleright B, \Delta) = \wedge[\Gamma] \rightarrow (B \vee \vee[\Delta])$

$T_{PL}(\Gamma \blacktriangleright (A \blacktriangleright B), \Delta) = \wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$

So we need to show that $\wedge[\Gamma] \rightarrow (B \vee \vee[\Delta]) \vdash \wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$

Proof (by Natural Deduction):

- | | | |
|-----|--|----------------------|
| (1) | $\wedge[\Gamma] \rightarrow (B \vee \vee[\Delta])$ | Premise |
| (2) | $\wedge[\Gamma]$ | Assumption |
| (3) | $B \vee \vee[\Delta]$ | 1,2 \rightarrow -E |

(4) B	Assumption
(5) A	Assumption
(6) $A \rightarrow B$	5,4 \rightarrow -I
(7) $(A \rightarrow B) \vee \vee[\Delta]$	6 \vee -I
(8) $\vee[\Delta]$	Assumption
(9) $(A \rightarrow B) \vee \vee[\Delta]$	8 \vee -I
(10) $(A \rightarrow B) \vee \vee[\Delta]$	3, 4, 7, 8, 9 \vee -E
(11) $\wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$	2, 10 \rightarrow -I

QED

REX: $\Gamma \blacktriangleright (A \blacktriangleright B), \Delta \gg \Gamma, A \blacktriangleright B, \Delta$

$T_{PL}(\Gamma \blacktriangleright (A \blacktriangleright B), \Delta) = \wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$

$T_{PL}(\Gamma, A \blacktriangleright B, \Delta) = (\wedge[\Gamma] \wedge A) \rightarrow (B \vee \vee[\Delta])$

So we need to show that $\wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta]) \vdash (\wedge[\Gamma] \wedge A) \rightarrow (B \vee \vee[\Delta])$

Proof (by Natural Deduction):

(1) $\wedge[\Gamma] \rightarrow ((A \rightarrow B) \vee \vee[\Delta])$	Premise
(2) $\wedge[\Gamma] \wedge A$	Assumption
(3) $\wedge[\Gamma]$	2 \wedge -E
(4) A	2 \wedge -E
(5) $(A \rightarrow B) \vee \vee[\Delta]$	3,1 \rightarrow -E
(6) $A \rightarrow B$	Assumption
(7) B	4,6 \rightarrow E
(8) $B \vee \vee[\Delta]$	8 \vee -I
(9) $\vee[\Delta]$	Assumption
(10) $B \vee \vee[\Delta]$	9 \vee -I

(11) $B \vee \vee[\Delta]$ 5, 6, 8, 9, 10 \vee -E

(12) $(\wedge[\Gamma] \wedge A) \rightarrow (B \vee \vee[\Delta])$ 2, 11 \rightarrow I

QED

That concludes the proofs – there’s nothing that one can derive in Sequentese that one can’t derive in classical logic, or vice versa.

V. Conclusion

What exactly has been shown? Overall, the ambition has been to show that a classical propositional language involving logical connectives can be plausibly interpreted in an expressivist manner, i.e. in terms of complex propositional attitudes. That interpretation has no role for attitudes to logically complex propositions. Ultimately, our thought about the world doesn’t need to involve anything like logical connectives: if we could think and talk in something like Sequentese, we’d be able to express what we can now express, and reason as we now reason. It would be a kind of dispensability result, analogous perhaps to the Field project in the philosophy of mathematics, which aimed to show that our scientific discourse doesn’t ultimately have to involve mathematical vocabulary⁸⁷.

Some might worry that this ‘result’ is illusory: it is not enough to show that PL can be translated in a connective-less language, if that language itself cannot be explained without the use of logical connectives. Indeed, if one looks at the interpretation of Sequentese in Section III.II.I., words like ‘and’, ‘or’, ‘if’ and ‘not’ are all over the place. So has anything really been gained?

I think this worry is misplaced. If we look at how the model-theoretic interpretation of the connectives of PL is introduced in a standard logic textbook, we’ll see that this happens in English (or whatever language the textbook is in) and that English connectives are used: it is said, for instance, that a wff of the form “ $A \wedge B$ ” is true if and only if A is true and B is true. The understanding of the logical

87. Field (1980).

connectives of PL thus relies on the understanding of the logical connectives of another language. But it needn't have been English. It could even have been PL itself, if that were the reader's language.

I've given the interpretation of Sequentese in English, in a language, that is, that uses connectives. But I needn't have. My readers *could* have spoken a connective-less language, like Sequentese. We know that there are such languages because Sequentese is one of them.⁸⁸ If that had been the way in which the interpretation of Sequentese had been presented, then the understanding of Sequentese would not have involved a prior understanding of some logical connectives. The fact that we presently speak a language that has connectives is not a deep metaphysical fact. It does not tell us anything about the structure of reality. If we could not have spoken a language without connectives, for some appropriately deep reason, that would show something. But this is not the case, as the example of Sequentese testifies. A language like Sequentese, combined with some non-logical terms describing the basic types of propositional attitudes and their relations, could be used to give an account of the logical connectives which did not presuppose an understanding of the logical connectives. And that is enough for the purpose of drawing metaphysical conclusions.⁸⁹

The dispensability of the logical connectives is an interesting result in its own right, even if one isn't particularly interested in the metaphysical upshot. It has a bearing on other matters in the philosophy of logic as well. Take, for instance, the epistemology of the basic inference rules. A popular view, associated in particular

88. Granted, one would have to do a little work to turn Sequentese into a usable language: one would have to specify what it is one does by saying or writing down a sequent of Sequentese. One proposal is to say, first off, that by uttering a sequent, one expresses the attitude that the sequent names (always a ruling out). Then one could add that, when the ruling-out in question leaves open only one pattern of attitudes towards the propositions involved in the sequent, it is conversationally implied that one instantiates that pattern. In this way one could use the language to express the acceptance or rejection of any proposition. To take a simple example: by uttering $\blacktriangleright P$, one expresses the ruling-out of the pattern that consists in rejecting P . That leaves open only the pattern of accepting P , which attitude is then implied. Alternatively, one could just enrich the language with (non-embeddable) devices signifying the acceptance or rejection of a proposition, and add some rules predicting the implicatures generated when these acceptances and rejections are uttered alongside sequents. For instance, if I utter $P \blacktriangleright Q$ and use another device to signal my acceptance of P , that might imply my acceptance of Q .

89. Cf. Melia (1995) for thoughts on how the mere possibility of a language can prove a metaphysical point.

with Paul Boghossian and Christopher Peacocke, says (in a nutshell) that our capacities to make basic inferences are to be explained in terms of our possessing basic logical concepts, corresponding to the logical constants.⁹⁰ We possess these concepts, and knowing which inferences they legitimise is part of what it is to have these concepts, so we do know which inferences are legitimate. However, if the logical constants (at least the logical connectives) are dispensable, then it seems one could be a classical reasoner without having the concept of, say, the conditional. That throws a spanner into the Boghossian/Peacocke story, and gives some support to the view that our possession of logical concepts is to be explained in terms of our capacities to reason logically, rather than vice versa. However, it is not my purpose in this thesis to pursue the debate on the epistemology of logic, interesting as it is. But it seems worth pointing out that the expressivist view interacts with these other debates in the philosophy of logic as well – and these interactions may be worth pursuing at some point.

Sequentese is not yet complete, for it only tackles the logical structure that is represented by the connectives. In chapter V, we will find ways of extending Sequentese to capture the structure we normally represent with the quantifiers. That extension will be based, roughly, on some ideas of P.F. Ramsey. Predicate Sequentese, as I will call the resultant language, goes beyond Propositional Sequentese in dispensing not only with logically complex propositions, but also with atomic propositions.

But first, in the next chapter, we take Propositional Sequentese and apply it to the problem that was left unsolved in Chapter I: giving an account of the semantic determinacy of the logical vocabulary. The solution will be, in rough outline, to allow the ideology of complex propositional attitudes, which the logical expressivist uses to interpret the logical connectives, to play a privileged role in the interpretative practices of a (Lewisian) radical interpreter. On this proposal, a radical interpreter is duty-bound to represent subjects as possessed of an array of complex propositional attitudes and consequently duty-bound to interpret their language as suited to the expression of complex propositional attitudes. This means

90. See, among others, Boghossian (2003) and Peacocke (1992).

that some parts of the language will have to be interpreted as playing a systematic role in forming linguistic vehicles for the expression of complex propositional attitudes. These parts of the language will come out behaving like the logical connectives.

Chapter IV: Radical Indeterminacy Revisited

0. Introduction	145
I. Chapter I Revisited	147
II. Metasemantics and Indeterminacy	149
II.I. A Lewisian Approach	149
II.II. Radical Interpretation	150
III. Lewis and Naturalness	155
IV. The Metasemantics of the Logical Connectives	158
IV.I. Psychological Assumptions	159
IV.II. Assumptions about Language	162
IV.III. Pinning Down the Constants	163
V. Reversing the Polarity on the Reference Magnets	165
VI. Conclusion	169

0. Introduction

In earlier chapters, we've discussed some arguments given by Ted Sider for the view that I've been calling logical realism. According to this view, the world comes equipped with a special kind of metaphysical structure, which I dubbed logico-metaphysical structure. This structure is supposed to be reflected in our first-order logical vocabulary.⁹¹ In terms of the now familiar metaphor, the connectives and quantifiers are thought to carve at some joints in reality: logical joints. We have reason to posit this kind of structure, Sider claims, because we need it to explain some things. One of the things that need explaining is why there is a logical

91. Sider (2009a).

vocabulary at all, and why it crops up in all of our theories. Another thing that needs explaining is how the terms of this vocabulary acquire determinate meanings, in the face of some well-known worries about the radical underdetermination of meaning by use.

Logico-metaphysical structure would indeed help us explain these things. We could explain why there are logical terms in our theories if we could point out something that these terms *get at*, some feature of reality that, if we did not have these terms, we would be failing to get to grips with. If there were some important feature of reality which our logical terms represented, then that would explain why our theories have those terms. Logico-metaphysical structure is, by definition, that sort of feature. Logico-metaphysical structure would also help us explain why the terms of the logical vocabulary have determinate meanings. It may be that the facts about how we use a symbol like “&” don’t determinately fix a semantic value for that symbol. But the term might still take on a determinate meaning if there were, among those candidate meanings which roughly make sense of the usage facts concerning “and”, some candidate which is more eligible than the others. If the world had logico-metaphysical structure, and if some part of this structure were the 'conjunctive' bit, then that bit of structure could be the reference magnet that is needed to give “&” a determinate meaning.

If a (healthy) metaphysical squeamishness stops us from wanting go along with Sider's metaphysics of logic, we need to address these arguments. In particular, we have to show how the explanatory tasks – which I regard as genuine – can be dealt with without appealing to logico-metaphysical structure. If there were good alternative explanations available, this would undercut the motivations that Sider has given us for believing in this sort of structure⁹².

The previous chapter, chapter III, was devoted to performing part of the first of these two explanatory tasks. In that chapter, we gave a story about the usefulness of the logical connectives; an explanation of what they help us to do. In brief, the story was that these terms help us to express a great variety of complex

92. Although it wouldn't, of course, show that the logico-metaphysical structure isn't there. But since logico-metaphysical structure is a theoretical posit, the burden of proof is on the positer.

propositional attitudes. We need to be able to express such attitudes if we want to use our language to exchange views, disagree, argue with each other, and convince each other of things. I take it as uncontroversial that activities like that are useful and important. And because they are, so are the logical connectives.

In this chapter we turn to the other explanatory task, the task of explaining how the terms in the logical vocabulary have determinate meanings. It will turn out that the logical expressivist story which we set out in the previous chapter will also help us to explain the semantic determinacy of the logical vocabulary.

I. Chapter I Revisited

Although more than one argument for radical semantic indeterminacy is around in the literature, in chapter I we concentrated on one argument, known as the Kripkenstein or Rule-Following problem.⁹³ In his (2009a) Sider mentions that a problem of the Kripkenstein kind can be raised for the logical connectives, although he does not spell it out. In chapter I, I did spell the problem out. For the case of the logical connectives, I defined a candidate meaning for the logical conjunction symbol “&” which, though obviously strange and gruesome, is consistent with the way that “&” is used by us⁹⁴. For the case of the quantifiers, I defined a candidate meaning for the universal quantifier “ \forall ”, which is similarly weird but consistent with the usage facts.

Sider, we saw, can explain why the meaning of “&” is not indeterminate between **and** (the 'intended' meaning of “&”) and **quand**, and all the other gruesome candidate meanings in the same ballpark as **quand**. It's because **and** is special: this particular function from truth values to truth values is a particularly *natural* one, and that's ultimately why “&” has that meaning, even though the facts about how we use “&” don't rule the other candidate meanings out. On Sider's Lewisian metasemantics, the specialness of **and** gets to play a role in the determination of

93. Kripke (1982).

94. And with our intentions and dispositions to use “&” in certain ways.

the meaning of “&”. **And** is a 'reference magnet' and **quand** is not.

Sider is in a position to say something similar about “ \forall ”. Here we need to be somewhat careful. I argued in chapter II that the particular externalist metasemantics for quantifiers that Sider proposes is inadvisable. But in that chapter I distinguished, with respect to quantifier meanings, two distinct meaning-fixing tasks. If the notion of a quantifier meaning is the 'local' one – as Sider's looks to be – then one task is fixing the domain-aspect of the quantifier: determining what the quantifier ranges over. The other task is fixing what kind of function the quantifier performs on that domain – in the case of the existential quantifier, mapping a property to the True iff its extension is one of the non-empty subsets of the domain. If the notion of a quantifier meaning is the 'global' one, there is only the latter metasemantic task, because that sort of quantifier meaning is not domain-specific.

The worries I articulated in chapter II were specific to the task of fixing the domain-aspect of quantifier meanings. But there is still the option of proposing, on separate grounds, an externalist metasemantics for what I described as the properly 'logical' aspect of quantifier meanings. As I showed in chapter I, there is a Kripkenstein problem to be formulated for quantifiers, and it concerns precisely that logical aspect of quantifier meanings. So just as Sider posits some reference magnets to solve the underdetermination problem for connectives, he can posit some reference magnets to solve it for the quantifiers. That would be a separate externalist commitment, not open to the particular challenge raised in chapter II. It seems likely that if Sider were persuaded of the semantic distinctions made in chapter II, he would make this separate commitment.

In chapter I, I conceded that a Lewis-style externalist metasemantics provides a plausible explanation of the determinacy of meaning in the face of Kripkenstein scepticism. Sider's response to the threat of Kripkenstein underdetermination in the logical vocabulary is an extension of that approach, and so it also looks promising, *prima facie*. But it seems to commit us to logico-metaphysical structure. I want to avoid that particular metaphysical commitment, so I can't say what Sider says.

In this chapter I will argue that – to daringly mix metaphors – we can have the

Lewisian cake without biting the Siderian bullet. We can concede that an appeal to reference magnetism provides a response to the Kripkenstein worry. But we can challenge the assumption that the mechanism of reference magnetism, when it operates, always involves positing some metaphysical structure. An alternative that I'll present is that certain terms get assigned certain meanings because our interpretative practice dictates that *some* terms in the language will have to be assigned those meanings, and the terms in question are the best candidates to bear those meanings, given the way we use them. I'll argue that this different form of reference magnetism is at work in the case of the logical connectives.

But before we can present that argument properly we need to have a proper metasemantic framework in place, to give us something to tinker with.

II. Metasemantics and Indeterminacy

II.I. A Lewisian Approach

Lewis wrote quite a bit on metasemantics, and his views developed over time, but there is a core set of ideas that remains in place throughout. One of the most significant adjustments in his view was the inclusion, from "New Work" (1983) onwards, of an explicit externalist element, to wit an appeal to a distinction between more and less natural properties. That distinction plays a number of roles in his philosophy, but in his metasemantics one of the most important roles it plays is to furnish a response to the sorts of indeterminacy worries that Putnam and Kripke raised.⁹⁵

The Lewisian metasemantics I'll set out is primarily derived from an earlier paper, "Radical Interpretation" (1974). In that paper, the task of describing how meaning facts get determined is pictured as the task of describing how an ideal radical interpreter would figure out the meaning facts. The interpreter is ideal, in having

95. Lewis (1974), (1975), (1983), (1984) and (1992).

access to facts that we don't (fully physical descriptions of the subject, their past, and their environment) and unlimited processing powers. But the interpretation is radical, in that the interpreter doesn't have any prior information about the subject that's couched in semantic or intentional vocabulary.

Lewis's metasemantics was described in an earlier chapter as *interpretationist*: the semantic facts, for Lewis, are those facts that the best interpretation of our linguistic behaviour (i.e. the best semantic theory) predicts. Lewis's interpretationism can be seen at work in "Radical Interpretation". Lewis takes it that by describing what the ideal radical interpreter would do – what facts they would use, what principles they would use to derive things from these facts – we end up describing not just how semantic facts are known or knowable, but how the semantic facts are determined by the non-semantic facts.⁹⁶ Ideal & radical epistemology of meaning equals metaphysics of meaning.

After setting out this Lewisian metasemantics, I'm going to see how we might add in the later-Lewisian appeal to naturalness. As it turns out, there are a number of ways of supplementing the initial story so that plausible results come out. Just saying that there is an objective distinction between the more and less natural won't be of much use: some mechanism needs to be indicated by which the relative naturalness of properties can get to bear on the determination of meaning. But once such a mechanism is in place, it will turn out that it can be adapted to uses that don't involve naturalness – other factors besides naturalness could potentially bear on meaning in comparable ways.

II.II. Radical Interpretation

Lewis sets out the challenge of radical interpretation in the following terms. We imagine we're trying to understand the thoughts and words of Karl. All that we know about Karl are the physical facts. We know what noises he makes in what situation, what the world he's in is like, what his brain states are like, and any other

96. Lewis (1974), p. 333.

physical facts that we might want to know. Call the description of these facts P. What we want to find out is what Karl thinks (more specifically what his beliefs and desires are) and what his utterances mean (the meanings of his sentences, the referents of his words and the compositional structure of his language). When we're trying to establish what Karl thinks, we're really after two distinct things: we want a description of Karl's beliefs and desires in our language (call this description A_o) and we want a description of these states in Karl's own language (call this description A_k). Finally, call the semantic description of Karl's language M.⁹⁷

A caveat: we might think that it's misguided to try and look at Karl in isolation, and try to determine his meanings on that basis. Karl is presumably a member of a linguistic community, and meaning (we might think) is produced by a community of language users, not by individuals. I think that's a point worth noting, but I think we can view Lewis's focus on Karl as a harmless idealisation. Nothing in the resulting story, as far as I can see, resists being extended to the radical interpretation of a whole community. Focusing on Karl simply makes things a little easier at a first pass, so I'll follow Lewis in just looking at Karl.

The radical interpretation challenge is to find out what A_o, A_k and M are, using just the data in P. The interpreter does this by means of a set of general principles which tell us how beliefs and desires are related to meaning, and how both of these are related to behaviour and environment. Lewis calls it a 'general theory of persons'. A general theory of persons, combined with appropriate inputs, can be used to generate a particular theory of a particular person: in our case, a 'theory of Karl', which tells us what Karl thinks, desires and means. This particular theory needn't be exactly an instantiation of the general theory – Karl's beliefs, desires and meanings might be related in somewhat idiosyncratic ways – but it will be roughly an instantiation of the general theory.⁹⁸

We thus embark on our interpretation of Karl assuming two things. One, he has a system of beliefs, desires, and meanings. In other words, there is a particular theory of Karl to be had. Two, this theory will fit with the general theory of persons,

97. Ibid. pp. 332-4.

98. Ibid. p. 334.

because the general theory of persons tells us what it is for a certain system to be a system of beliefs, desires and meanings. These notions are implicitly defined by the general theory.⁹⁹

Lewis then gives us some principles which he thinks are part of our general theory of persons:

1. The Principle of Charity. This principle allows Ao to be constrained by P. It tells us that Karl's beliefs and desires are as they ought to be, given his history. They ought to make sense, given what he has experienced and has been exposed to. One way of glossing that, which Lewis gives us, is that we're supposed to ascribe to Karl the beliefs and desires that we would have, if we had his history of experiences. That might mean, in a lot of cases, that we ascribe to him beliefs that we think are true, and desires for things we think desirable. But it might mean in other cases that we ascribe beliefs to him that we think are false, if we wouldn't expect ourselves to have believed the truth, given Karl's evidence.¹⁰⁰
2. The Rationalisation Principle. This principle likewise allows Ao to be constrained by P. It tells us that Karl's behaviour (described so far in non-intentional vocabulary) ought to make sense, rationally, given his beliefs and desires. If Karl performs some action, then that action ought to come out as a good idea, given his beliefs and desires. Lewis thinks we could bring decision theory to bear on this matter: the complete set of Karl's actions ought to come out as the one with maximum expected utility, given his beliefs and desires.¹⁰¹
3. The Principle of Truthfulness. This principle constrains Ao and M. Those of Karl's beliefs and desires that concern his linguistic behaviour ought to conform to a 'convention of truthfulness in Karl's language'. This

99. Ibid. p. 335.

100. Ibid. pp. 336-7. Even under conditions of idealisation, it might be a tall order to fill in all of Karl's beliefs on this basis. Even if we know that Karl has, say, read *Moby-Dick*, there might not be a fact of the matter about what beliefs he then ought to have about the book. But for lots of basic, perceptual beliefs, it is plausible that we could fill in Karl's attitudes in this way, given the facts about his history of evidence.

101. Ibid. pp. 337-8.

means that when some utterance U1 means that there's an apple on the table, then Karl ought to have (i) a desire not to utter U1 unless there's an apple on the table; (ii) a belief that others in his linguistic community have such a desire; (iii) a belief that there's an apple on the table, when someone utters U1; (iv) a belief that others form that same belief in the same situation; (v) a belief that others expect him to have attitudes (i)-(iv); (vi) a belief that they expect him to have belief (v). Lewis leaves open that there might be more constraints, in the same spirit.¹⁰²

4. The Principle of Generativity. This constrains M. The theory of meaning that we give for Karl's language (i.e. M) should assign truth conditions to his sentences in a way that's tractable: it has to be finitely specifiable, and it may also have to be simple in uniform to some extent as well. After all, we want Karl's language to be a language that Karl could understand.¹⁰³
5. The Manifestation Principle. This constrains Ak and M, and to some extent Ao. Unless we have reason (on the basis of how we've filled in Ao) to ascribe to Karl some reason for deception, the sentences Karl utters ought to express propositions that he believes. Lewis also thinks there should be a similar principle for desires, though he claims not to know how it would run.¹⁰⁴
6. The Triangle Principle. This constrains Ao, M and Ak all at once. It says that Karl's attitudes should be the same, whether described in our language or in Karl's own language. So M ought to be such that when we ascribe Karl a certain attitude A (in our language, so as part of Ao), we also get to ascribe Karl the attitude that results from translating A using M.¹⁰⁵

Lewis then describes several procedures for using these principles to get Ao, Ak and

102. Ibid. pp. 338-9.

103. Ibid. p. 339.

104. Ibid. p. 339.

105. Ibid. p. 339.

M from P. The procedure he favours has the following steps¹⁰⁶:

- 5) Using the facts about Karl's behaviour and the facts about his history of evidence (both sourced from P) fill in Ao using the Principle of Charity and the Principle of Rationalisation.
- 6) Using those facts from Ao (as tentatively filled in) that pertain to speech behaviour, fill in M so that, first off, the Principle of Truthfulness is satisfied and, second, the Principle of Generativity.
- 7) Using the now filled in Ao and M, fill in Ak by means of the Triangle Principle. This should automatically satisfy the Manifestation Principle as well.

Lewis notes that if we're really taking the interpreter to be ideal, they wouldn't strictly speaking need to follow any stepwise procedure to settle the facts about Karl: they could just try out all possible complete solutions, and see which one fits the constraints the best.¹⁰⁷ Nevertheless, such a step-wise procedure may make it more clear how it is that the six principles above enable P to determine Ao, M and Ak.

What hasn't been settled by any of the above is how determinate the facts in Ao, M and Ak end up being. Lewis indeed admits he hasn't done much to show that the ideal interpreter will be able to settle things one way or another.¹⁰⁸ He notes three ways in which indeterminacy could arise. First, the truth conditions of the full sentences in M may not determine the meanings of the constituents of those sentences. Lewis thinks it's likely that they won't, considering cases like Quine's 'gavagai'.¹⁰⁹ Second, there may not be any way of satisfying all the metasemantic constraints listed in Karl's case, so that the interpreter would have to somehow pick the best of a bunch of compromise solutions. There may not be a fact of the matter about what they would choose.¹¹⁰ It's hard to figure out how worried Lewis is about these two sorts of indeterminacy – all he does is grant that they exist, and

106. Ibid. pp. 341-2.

107. Ibid. p. 342.

108. Ibid. pp. 342-3.

109. Ibid. pp. 342-3.

110. Ibid. p. 343.

then set them aside. There's a third sort of possible indeterminacy which it seems might worry him more: the possibility of multiple different solutions fitting all of the metasemantic constraints perfectly. With regard to this sort of indeterminacy, his 'credo' is that if it could be shown to arise, on the constraints that he has outlined, the conclusion he would draw is that there must be further constraints. In other words, this sort of indeterminacy is not to be granted.¹¹¹

III. Lewis and Naturalness

As we just saw, in his (1974) Lewis is aware of the abstract possibility of excessive semantic indeterminacy and takes a very general stand on it: a metasemantics which generates that sort of indeterminacy needs to be tinkered with until it doesn't. From his (1983) onwards, this issue seems to have worried him enough so that he actually does propose further metasemantic constraints.

As we explained in Chapter I, Lewis's preferred solution employs an ideologically primitive distinction between an 'elite' of perfectly *natural* properties and a multitude (or possibly 'proletariat') of non-perfectly-natural properties. Natural properties are glossed by Lewis as those which make for objective similarity between things: qualitative duplicates would be objects that share all their natural properties.¹¹² Epistemically, Lewis thinks of the perfectly natural properties as those that would be talked about by a hypothetical perfected fundamental physics.¹¹³

The distinction between natural and non-perfectly-natural properties is not one of degree, but it can be used to create a distinction of degree. For instance, we could take a language in which we only have predicates expressing natural properties, and then define the less-than-perfectly-natural properties in them.

111. Ibid. p. 343.

112. Lewis (1983), pp. 355-6. That's how Lewis puts it, but it's slightly ambiguous. Really, they are qualitative duplicates if their parts can be put into one-one spatial correspondence and all of those parts share all of their natural properties with their correspondents. The natural/unnatural distinction also plays a variety of related metaphysical jobs. In virtue of playing a role in defining intrinsic duplication, naturalness also has a role to play in explaining the notions of supervenience (for which we need a notion of duplicate worlds), and divergence between worlds.

113. Ibid., pp.356-8.

Depending on how complicated the definition of a given less-than-perfectly-natural property ends up being in that language, we can assign it a complexity value (perhaps on the basis of how many logical connectives the definition contains) and this complexity value could serve as a measure of *relative* naturalness. So the perfectly natural properties would have minimal complexity values, and would therefore be the most relatively natural properties, and other properties would be less natural the more complex their definitions.¹¹⁴

A distinction between more and less natural properties can help us combat various kinds of indeterminacy that might arise on the above metasemantic picture. First, there might be indeterminacy in Ao: when we try to satisfy the Principles of Charity and Rationalisation, there can be perverse candidate assignments of propositional attitudes that satisfy these constraints. On the basis of Karl's verbal and non-verbal behaviour around emeralds and other things that look to have the same colour as emeralds, we could plausibly assign to him the belief that emeralds are green; but the belief that emeralds are grue is also a candidate, and it might satisfy the Principles, as stated. But we might try and build into those principles some kind of bias towards the more natural properties: we could say that an assignment of mental states to Karl which involves more natural rather than less natural properties (so green, rather than grue) is more charitable, and makes him more rational.

Second, there might be indeterminacy in M: if by hypothesis we can arrive at a single acceptable assignment of truth conditions to Karl's sentences, there will be multiple ways of giving meanings to the parts of Karl's sentences which satisfy the Principle of Generativity. We might know, when Karl's sentence "Emeralds are grrr" would be true. That doesn't tell us whether "grrr" should be interpreted to denote the colour green or the colour grue. But again, we could build into the Principle of Generativity some bias towards the more natural properties: we could say that an assignment of denotations to Karl's words which assigns to his predicates more rather than less natural properties is a better satisfier of the Principle of

114. See Williams (unpublished) for suggestions on how the details of this might (or might not) work out.

Generativity.

In this sort of way various kinds of indeterminacy might be combated. Lewis himself does not, in the papers in which he writes about naturalness, go into a huge amount of detail about how exactly he sees relative naturalness playing a role in his metasemantics. As a result there is some controversy about the role it plays, and the importance Lewis accorded it.¹¹⁵

Williams (draft) has put forward an attractive proposal for how we might think of naturalness bearing on meaning in Lewisian interpretationism, which would in principle allow it to perform a number of tasks of the above sort. Williams understands the role of naturalness in metasemantics as an instance of the role that naturalness plays in Lewis's theory of natural laws. Lewis has a 'best theory' view of natural laws: of all the theories that make sense of the regularities among physical events in the world, one is the best. Whatever laws of nature that theory encodes are the actual laws of nature. Naturalness plays a role in Lewis's view natural laws because it helps constrain the best-ness of theories. Theories are better, *ceteris paribus*, if they are simpler.

But what is simplicity? Problematically, the same theory can appear simpler or more complex, depending on how it is written down. A theory which consists in the claim that emeralds are green is simpler than a theory which says they are grue, when these theories are written down in a language which has "green" as an undefined term, but not "grue". But the reverse would be the case with a language which had "grue" as an undefined term, but not "green". So what is the language we should be writing theories down in, if we want to get an objective measure of simplicity? Lewis's answer is: a language in which the undefined predicates pick out perfectly natural properties.

We can apply this story to metasemantics. Lewis's metasemantics is, as we've said, interpretationist. It is, like Lewis's theory of laws, a theory on which the facts concerning a certain subject are the ones projected by the best theory of that subject. So we can allow naturalness to bear on meaning by allowing it to bear on

115. See for instance Schwartz (unpublished) and Weatherson (unpublished).

the comparative goodness of semantic theories. Semantic theories, like any theories, are better if they're simpler. Simplicity we can understand as simplicity-in-a-canonical-language, where the canonical language is one in which the undefined predicates denote perfectly natural properties.

This looks to predict the right results. Take our total theory of Karl, i.e. the theory which fills in Ak, Ao and M. This theory will be simpler if the bit of it which fills in Ao is simpler: if, for instance, it describes Karl's attitudes to emeralds in terms of greenness rather than grueness. (Green, I assume, is more relatively natural property than grue.) Our theory of Karl is also simpler if it fills in M in a simpler way, and one way of doing that is by having Karl's word "grrr" denote greenness rather than grueness.

This means that we wouldn't need to build in an appeal to naturalness in the various Principles on a case-by-case basis. By letting naturalness factor into theoretical simplicity directly, we can have it play a role in the application of all of these Principles.

IV. The Metasemantics of the Logical Connectives

We now have a basic Lewisian metasemantics in place, one which looks like it can deal with threats of radical indeterminacy, by means of a general appeal to naturalness. Let us now return to the matter in which we're specifically interested: the meanings of the logical connectives. These, we've seen, also seem threatened by radical indeterminacy. But we've been trying to get away with not assuming any logico-metaphysical structure. Sticking to that resolve means that our hypothetical canonical language, in which one measures simplicity, can't have logical connectives which denote perfectly natural logical operations. There are no perfectly natural logical operations, in our metaphysical picture.

That means that an interpretation of a language which interprets its connective "&" as meaning **quand** rather than **and**, wouldn't come out as less simple. This is because the canonical language in which we measure simplicity won't have any

opinion on which meanings for “&” are simpler. Since we do think that “&” has a determinate meaning, how is this to be explained?

I think that the Lewisian metasemantics that we've set out above has surprising resources for giving a story about the meaning of the logical connectives, when we combine it with the logical expressivism set out in the previous chapter. Let me draw attention to some important features of this metasemantics.

IV.I. Psychological Assumptions

In Lewis's metasemantics, it is assumed that the language to be interpreted is used by the interpretees to communicate their beliefs and desires. It is a guiding thought of the theory that words are meaningful in virtue of being used by people to communicate their beliefs and desires. This is a very natural thought: after all, that's what we think language is for. That means it is assumed from the get-go that that people *have* beliefs and desires. If they did not have such attitudes, then the sounds they made would not be proper objects of semantic interpretation, in the same way that rustlings of leaves in the forest are not fit subject for interpretation.

For my purposes the presence of desires in the story not so relevant – I'm going to set that aside – but the presence of beliefs is very important. The orthodoxy has it that people have beliefs in both simple and complex propositions. Lewis believes this as well. The logical expressivist disagrees with that. There are beliefs in simple propositions, but not beliefs in complex propositions. Instead, there are complex attitudes that involve, as their content multiple simple propositions.

We can adapt Lewis's metasemantics to incorporate this aspect of logical expressivism. Instead of the assumption that Karl has beliefs, we're going to have the assumption that Karl accepts some simple propositions, rejects some simple propositions, and has complex attitudes involving multiple propositions. In other words, we attribute to him a set of attitudes that is as rich as the set of beliefs that the Lewisian ideal interpreter attributes to him. But it is a more varied set of attitudes: it feature all the sorts of attitudes that were argued for in chapter III. For

example, where the Lewisian radical interpreter would attribute to Karl a belief that it's either cloudy or sunny, the expressivist radical interpreter attributes to Karl the ruling out of the pattern of attitudes that consists in rejecting that it is cloudy and rejecting that it is sunny. And so on.

If the interpreter assumes that Karl has these sorts of attitudes, the task of filling in Ao is the task of filling in these attitudes for Karl. If it is indeed possible to fill in Ao on the original Lewisian story, then it seems it will also be possible, derivatively, to fill in Ao on this revised story. For, as we've shown in the last chapter, we can provide a Sequentese translation of any logically complex sentence. Wherever the Lewisian radical interpreter says "Karl believes that [logically complex sentence S]" the expressivist can say "Karl has attitude $[T_{seq}(S)]$ ". The Principle of Charity that is used to fill in Ao will have to be slightly different, though. We're not simply trying to maximise the number of rationally justifiable beliefs Karl has, but more broadly the number of rationally justified propositional attitudes, where these can be acceptances, rejections, and all manner of complex attitudes. But again, if we can do the former, we can do the latter.

It will be somewhat roundabout, though, to first assign Karl a bunch of beliefs, in the way described by Lewis, and then replace those assignments with assignments of complex propositional attitudes. There will be a more direct way of doing this, by reformulating the principles that Lewis uses to constrain Ao, i.e. the principles of Charity and Rationalisation. These principles basically say that we should attribute to Karl the beliefs that we should expect him to have, given his evidence and behaviour, and assuming his rationality. That means we should attribute to him true beliefs about things that he ought to be expected, given his situation, to be aware of, and false beliefs about things that he ought to be expected to be deceived about.

We hadn't assumed in chapter III that complex propositional attitudes are the sorts of things that can be true or false, and it is not immediately how truth and falsity would apply to them. Usually we think of the truth or falsity of beliefs as deriving directly from their contents. A belief is true iff the proposition that is its content and false otherwise. But since complex propositional attitudes can have multiple propositions, the relations between the truth values of these propositions

and the attitude in question would have to be more complex. Ultimately, I think “true” and “false” are not entirely natural-sounding terms in this context, so I will instead introduce a more generic notion of *correctness* for propositional attitudes.

Correctness, as a predicate of attitudes, can be defined inductively as follows. Note that attitudes of ruling in and ruling out can be divided into levels, according to the levels of the patterns that they rule in or rule out (see chapter III for the division of patterns into levels. Then:

Correctness 0-a: An attitude of accepting a proposition P is correct iff P is true;

Correctness 0-b: An attitude of rejecting a proposition P is correct iff P is not true;

Correctness 1-a: The ruling-in of a level-1 pattern is correct iff all of the attitudes in the pattern are correct;

Correctness 1-b: The ruling-out of a level-1 pattern is correct iff some of the attitudes in the pattern are not correct;

Correctness 2-a: The ruling-in of a level-2 pattern is correct iff all of the attitudes in the pattern are correct;

Correctness 2-b: The ruling-out of a level-2 pattern is correct iff some of the attitudes in the pattern are not correct;

...

...

Correctness N-a: The ruling-in of a level-n pattern is correct iff all of the attitudes in the pattern are correct;¹¹⁶

Correctness N-b: The ruling-out of a level-n pattern is correct iff some of the attitudes in the pattern are not correct;

With these definitions, we can say, of a given complex attitude, exactly what it would take for it to be correct. We can then reformulate the principles of Charity

116. Note that by the definitions of the different levels of patterns, the attitudes of the patterns will always be of a lower level, and so their correctness conditions will have been defined at that point.

and Rationalisation so that they tell us to attribute correct complex attitudes about a certain matter to Karl when we have reason to expect him, on the basis of his evidence and behaviour and the assumption of his rationality, to be aware of how stand with that matter, and incorrect complex attitudes when we have reason to expect him to be mistaken.

Using those adapted principles and the information in P, we can then attribute a set of complex attitudes to Karl. Note that, by doing so, we would be attributing to Karl a set of more fine-grained things that we would on Lewis's story. Beliefs only have as much structure as the propositions that are their contents, and Lewis is intensional about propositions: they are just sets of worlds.¹¹⁷ The propositional attitudes we are attributing are, by contrast, structured entities. There seem to be two options at this point. When we have assigned a certain correctness condition to Karl, we can either (a) attribute to him all of the attitudes that have that correctness condition or (b) indeterminately attribute at least one of those to him. I think the latter would be preferable, since I don't want to blithely assume that the attitudes are implemented in Karl's head in such a way that it would be possible for him to have endless numbers of them. But for our current purposes, it won't matter a great deal what we do. Later on in the story, we can narrow down Karl's attitudes some more.

IV.II. Assumptions about Language

A second basic assumption in Lewis's metasemantics is that sentences have the function of expressing beliefs (and desires). The filling-in of M must therefore meet some general constraint. Among the sentences that can be formulated in Karl's language, some must be able to express beliefs. It can't be the case, for instance, that Karl's language comes out as not having anything like a declarative mood. That's not to say that there could not in principle be languages without anything like a declarative mood; languages in which one can only ask questions, for instance. But

117. Or sets of centred worlds, more precisely. But that's a complication we can ignore.

Lewis's metasemantic theory, given the way it is set up, does not extend to such languages. It is a metasemantics for languages that are used for communication between believers.

We can also adapt this aspect of Lewis's theory to expressivism. Instead of building in the basic assumption that sentences of Karl's language must be able to express beliefs, we build in the assumption that sentences of Karl's language must be able to express a variety of complex propositional attitudes as described in chapter III (and further in chapter V). That puts some constraints on how M can be filled in. Somehow, it must be possible, in Karl's language, to tell from a sentence what kind of propositional attitude it expresses. That will turn out to be important in a little bit.

The principles that Lewis proposes constrain M in two different ways. The principles of Truthfulness and Manifestation help us fill in M in a certain respect: they will allow us to assign truth-conditions to Karl's sentences. The Principle of Generativity plays a different role. Given an assignment of truth-conditions to Karl's sentences, that principle allows us to assign semantic values to the sub-sentential bits of Karl's language. The distinction between these two tasks is important for our purposes.

In the expressivist version of the Lewisian story, the first stage of filling in M will not be to assign truth-conditions to Karl's sentences, but correctness conditions. If we're assuming that the sentences of Karl's language are expressing beliefs, then it is appropriate to look for truth-conditions to assign them, for beliefs are the sorts of things that are truth-apt. Since we're assuming here that Karl's sentences express complex propositional attitudes that can be correct or incorrect, it would be appropriate to assign correctness conditions, rather than truth conditions, to Karl's sentences. The general idea would be that sentences will have the correctness condition of the attitude it is conventionally used to express.

IV.III. Pinning Down the Constants

With that notion in place, we can proceed to explain how M gets filled in, on the expressivist version of Lewis's metasemantics. We use a suitably tweaked version of the Principle of Truthfulness for this purpose: a Principle of (pardon my French) Correctfulness. The first step is to assign correctness conditions to Karl's sentences: we now know, of each of his sentences, that it is a vehicle for the expression of one of the attitudes with that correctness condition.

Then the Principle of Generativity comes in. We know that each of Karl's sentences must express some attitude that has the correctness condition assigned to that sentence. We also know that the structure of that attitude must be encoded, somehow, in the structure of that sentence. Here's why: since we're assuming that Karl's language is a device for communicating complex attitudes, it must be possible for sentences to somehow convey the structure of an attitude. The only plausible way in which this can be done that respects the Principle of Generativity is for the structure of the sentence to convey the structure of the attitude.

Even before we've pinned down the grammar of Karl's language, it ought to be possible to gauge the grammatical complexity that a given sentence could have (based on number of words, length of words etc.). This will allow us to narrow down the complexity of the attitude it might express, for the complexity of the attitude expressed cannot be greater than the grammatical complexity of the sentence expressing it (the grammatical complexity of the sentence might be much greater than the complexity of the attitude expressing it, though, because grammatical structure might be conveying some non-logical complexity in the content of the sentence).

We solve for M by trying out all the candidate grammars and vocabulary that meet certain conditions. They must have only a certain degree of complexity, given that Karl is only human. Likewise the vocabulary can only be of a certain size. The meanings assigned to names and predicates will have to be as eligible as can be managed. But last but not least, the grammar must assign, in a systematic way, to each sentence an 'expressive structure' that maps onto the structure of one of the attitudes that has the same correctness condition as that sentence. This is the structure that guarantees that the language can be used to convey specific complex

attitudes (which we are assuming it must be able to do).

Here's where the logical constants enter the picture. How might expressive structure conceivably be implemented in the language? There are many different devices that could play a role here: word order, conjugations/declensions of words, intonation. But it is very natural to think that vocabulary might play a role here: that the presence of certain words, and their place in the sentence, helps to indicate the expressive structure of the sentence. Using vocabulary for this purpose has something important going for it: it would allow one to stick arbitrary amounts of expressive structure in a sentence. It would serve better, therefore, to satisfy the demand of compositionality.

So the interpreter would be looking, in effect, for words in the language that can be attributed this role of helping to indicate expressive structure. They would be looking to identify a set of words that are distributed over the sentences of Karl's language in such a way that they serve to link up, systematically, Karl's sentences with attitudes that share the correctness conditions of those sentences.¹¹⁸ That set of words, I submit, is the set of logical constants. Given the role that these words play in generating expressions for attitudes of particular structures, it should then be possible to define these words in terms of the basic building blocks of those attitudes: acceptance, rejection, ruling out, etc. That is in effect what we've done in chapter III.

Importantly, we won't have to find anything out in the world to serve as the semantic values of such a word. A fortiori, we won't have to find anything relatively natural out in the world to serve as the semantic values of such a word.

V. Reversing the Polarity on the Reference Magnets

How does all this connect back to the issue of radical indeterminacy? Since on this story the interpreter doesn't assign any item of the furniture of reality as a meaning

118. Leaving open the possibility that some words will be identified as playing such a role in some circumstances, and perhaps not in others.

to the logical connectives, it can't be the naturalness of their semantic values that prevents the connectives from being radically semantically indeterminate. Nevertheless, their meanings are rendered determinate, though in an entirely different way.

There is a certain kind of 'reference magnetism' that happens in the case of the logical connectives. It is of a different sort than the reference magnetism that naturalness can create. When reference magnetism is due to naturalness, what happens is that the interpreter looks at the usage facts concerning a language (i.e. what Karl and his language-mates say) and assigns that set of meanings to the words of that language that, of those that satisfy the various principles, maximises naturalness. Here the usage facts, together with the Principles, can be viewed as generating a range of candidate meaning-assignments: all the meaning-assignments that satisfy the Principles, given the usage facts. Then from this range the naturalness-maximising candidate is picked.

In the case of the logical connectives, the order of the proceedings is reversed, in a way. We already have the interpretations sitting on a shelf: some of the words in the language will have to be interpreted as playing the merely expressive, structural role of the logical constants.¹¹⁹ Then, looking at the words of Karl's language and the correctness conditions of his sentences, we find which terms best fit this role.

This mechanism – call it 'reverse reference magnetism' – rules out the interpretation of “&” in our language as **quand**. “&-meaning-**quand**” is *not* a merely expressive device: given how it is defined, it interacts with the content of its 'quonjuncts' in a way that **and** does not. What a sentence that has “&-meaning-**quand**” in it expresses depends on what the sub-sentences are that it connects: it matters whether one of those sub-sentences expresses mystery proposition M or no. Since the usage facts about “&” are such that it is a prime candidate to be interpreted as a merely expressive device, it will get shanghaied into that role even if those usage facts are also consistent with it meaning **quand**.

This may seem a bit quick. Here are some qualms one might have.

119. On some occasions of use, at least.

First, one might think that “&-meaning-**quand**” can still, in a pinch, serve as a largely expressive device, if not as a *merely* expressive device. Might that not be good enough? I concede that I might. But what reason could the interpreter have for interpreting “&” in this way? By hypothesis, the **quand**-interpretation makes equal sense of the usage facts. It has no advantage over the **and**-interpretation in that regard. But the **and**-interpretation does have something over the **quand**-interpretation: it makes “&” a merely expressive device rather than just a largely expressive device. And by that token, an interpretation of Karl's language using **and** will make that language come out as more suited for the purpose which we have assumed it has: communicating propositional attitudes of varying complexity.

Second, how about the possibility that some interpretee happens not to have the usual repertoire of propositional attitudes? What if they have the attitude of rejection* rather than just plain rejection, where rejection* is somehow very subtly different from rejection in a way that's hard to detect? In that case, shouldn't the negation sign, in this interpretee's mouth, be interpreted differently? Suppose then that the facts about the interpretees are not rich enough to rule out that they have rejection* rather than rejection in their repertoire (let's say we can define rejection* in such a way that this comes out true). Wouldn't that mean that the meaning of the negation sign is indeterminate between expressing rejection and expressing rejection*, and moreover indeterminate between those and further subtle varieties on the usual attitudes?

Whether a real threat of radical indeterminacy looms here depends, first, on whether the scenario is possible. I don't know which construals of rejection* could be thought up, but it's not clear that analogues of grue or quus cases will really be available. Suppose that rejection* is an attitude that is much like rejection, but it behaves like acceptance during the first minute of every full moon. I don't think that any person could unambiguously be described as a rejecter* in this sense. For how isn't this 'rejecter*' just a person who accepts during the first minute of every full moon what they otherwise reject? I'm not convinced that we can gerrymander propositional attitudes in the same way as we can gerrymander meanings.

But if we can gerrymander propositional attitudes, it's a result that's neither here

nor there. It is built into the Lewisian metasemantics (or my variant on them) that we interpret the utterances of people by treating them as expressions of acceptances, rejections and complex propositional attitudes. Anyone who isn't interpretable as having these attitudes will simply not be a fit subject for interpretation. And if they are interpretable as having these attitudes, then that is how we will interpret them.

Of course, if there are alternative sets of propositional attitudes like rejection*, there may be possible interpretative practices which rely on interpreting people as having these alternative attitudes. But these practices aren't our practices, and unless we can be convinced that they are somehow better, it will presumably stay that way. I do not see this theoretical possibility as generating any meaningful indeterminacy. The meanings of the words in a language are determinate, by Lewisian lights, if there is one semantic theory such that it is the best theory for that language. The claim is not that the meanings of the words in a language are determinate if there is one metasemantic theory such that it is the best metasemantic theory and that according to this metasemantic theory, there is one semantic theory such that it is the best theory for that language. If we went that way, where would it end?¹²⁰

However, if one is not convinced by those considerations, there is more one could do to back it up. One could bring naturalness back into the story, in a different way. One could posit that the attitudes of rejection, acceptance, ruling in and ruling out are relatively natural, compared to rejection* and the like. They could be regarded as psychological natural kinds, of a kind discovered, however, by philosophical psychology rather than empirical psychology. That would not, I think, compromise the stated aims of this thesis to a significant degree. I aim was to avoid logico-metaphysical structure, which we identified, on the basis of what Sider says, as a novel sort of fundamental structure, not to be identified with something we already countenance. It is the fundamentality of this sort of structure, in addition to its being relatively ill-understood, that makes it objectionable.

120. Cf. Field (1975) for similar considerations on a slightly different topic (the role of causal relations in externalist metasemantics).

VI. Conclusion

In this chapter I've tried to bring some threads of this thesis together: the theory of the logical connectives that I offered in chapter III, which was meant in the first place to explain the presence of the logical connectives in our languages, turns out also to be relevant to another issue we've discussed earlier in the thesis, namely the spectre of radical Kripkenstein indeterminacy for the logical connectives. I've offered a story about how the meanings of logical connectives are determined that makes them into an interesting special case, different from the rest of the vocabulary. Some may see this as a disadvantage – the idea being that uniform accounts are better. But they're only *ceteris paribus* better, and here I think we have a legitimate exception. The terms of the logical vocabulary perform a different function than other terms do, and this justifies treating them as a special case, metasemantically.

There is still a gap in the overall story at this point. The expressivist story hasn't been extended to quantifiers yet. Since these are also part of the first-order logical vocabulary that this thesis focuses on, that might be considered odd. The reason they haven't come up in this chapter is that their treatment raises complications that were better postponed. But it is to this task – extending our account of the role and meaning of the logical vocabulary to the quantifiers – that the next chapter is devoted.

Chapter V: Expressivism for Quantifiers

0. Introduction	170
I. Strategies	172
I.I. Tractarian Quantification	173
I.II. Ramsey-Mill Quantification	175
I.II.I. Predication	176
I.II.II. Quantification	179
I.III. Correctness	181
I.IV. On the Way to Predicate Sequentese	183
II. Sequentese with Predicates: Syntax	185
II.I. Predication	185
II.II. Open and Closed Sequents	186
III. Sequentese with Predicates: Translation Manual	188
IV.I Sequentese with Predicates: Inference Rules	190
IV.II. Justifying the Inference Rules	197
V. Sequentese with Predicates is Standard Predicate Logic	198

0. Introduction

In chapter III, an ‘expressivist’ semantic theory was proposed for the logical connectives. The guiding idea was that the logical connectives allow us to form expressions that, rather than stating beliefs in logically complex propositions, express complex propositional attitudes to simple propositions. Logical complexity,

on this view, is to be found in sentences and in propositional attitudes, but not in propositions and not in facts. To utter a sentence of the form “ $A \vee B$ ” is not to express a belief in a proposition of a disjunctive form; it is to express that one is unwilling to both reject A and reject B. To utter a sentence of the form “ $\neg A$ ” is not to express a belief in a negative proposition: it is to express that one is unwilling to accept A.¹²¹ And so on and so forth.

I explained the proposal using an artificial language I dubbed ‘Sequentese’. Sentences of Sequentese do not express assertions: they express complex attitudes that involve multiple propositions, and what they express about these propositions is that the speaker is unwilling to adopt certain combinations of attitudes towards them. Attitudes toward single propositions (acceptances or rejections) can also be expressed in Sequentese, but as limit cases of complex attitudes.¹²² I explained how translations between Sequentese and the language of first-order propositional logic were supposed to go, and how one could systematically read off attitudes expressed from Sequentese expressions. Then I went on to show how reasoning in Sequentese looked, and that the Sequentese proof system, at least in the particular way I develop it, is a classical system.

The expressivist view of the logical vocabulary is geared toward meeting two theoretical challenges. As explained in chapter III, it explains the role of the logical connectives, and thereby explains why we should expect them to be in our vocabulary. This provides an alternative (a better alternative, hopefully) to explaining the presence of logical terms in our vocabulary in more robustly metaphysical terms, i.e. as a reflection of logico-metaphysical structure in reality. As explained in chapter IV, it also provides the materials for a metasemantics for logical terms that meets the challenge, articulated in chapter I, of staving off radical semantic indeterminacy. By securing semantic determinacy in that way, we would not need to do so by appealing to logico-metaphysical structure.

121. Note that the ‘unwillingness’ that is expressed is not just any old unwillingness: see chapter III for more details.

122. I did however leave open the option of extending Sequentese to include vehicles for the direct expression of acceptances and rejections.

But so far we've said nothing about the rest of the first-order logical vocabulary, viz. the quantifiers. As we saw in chapter II, it has been suggested that the quantifiers are metaphysically 'joint-carving' for reasons that don't generalise to the connectives (because they play a role in meta-ontology that the connectives don't). We saw already that those reasons were not conclusive, but one might nevertheless still be tempted to think, for other independent reasons, that the quantifiers are metaphysically significant in a way that the connectives aren't. Even if one did agree that the connectives played a purely expressive role, it wouldn't be crazy to regard the quantifiers differently. That means that we can't rest our case for the whole first-order logical vocabulary being metaphysically non-committing on the account we've given of the connectives. And so that's what this chapter is for: to show that there is a distinctively expressivist treatment of the quantifiers to be given.¹²³

I. Strategies

There are at least two basic routes open for extending the expressivist semantics to quantifiers. One would begin by reducing quantified talk to unquantified talk that employs only the connectives, and then treat those connectives in the way we've already explained. The other would begin by admitting that quantifiers bring something new to the table – they allow us express things that the connectives alone cannot – and try and specify what it is they express, i.e. describe the kinds of attitudes that quantified talk is the expressive vehicle for. I'll first talk a little about the first route; I don't think that it is ultimately the best way to go, but it's worth exploring the reasons why.

123. Of course, even if we did think that the quantificational vocabulary had a metaphysical import that the connectives didn't, there would still have to be a story about how the expressivist semantics for the logical connectives interacts with whatever non-expressivist semantics we would have for the quantifiers. If it turned out that there was no way to extend Sequentese to the quantificational vocabulary – not just no metaphysically non-committing way, but no way at all – then the logical expressivist project would seem to be doomed. Fortunately, things won't turn out so badly.

I.I. Tractarian Quantification

The easiest way one might extend Sequentese to the quantificational vocabulary is to try and get away with not adding anything to it. One could try and argue that Sequentese, as it is, is already expressively adequate, because the quantificational vocabulary doesn't add anything genuinely novel to the propositional fragment of the language – anything quantifiers can do, connectives, plus the non-logical vocabulary, can do already.

There is some precedent for this approach. It is most often associated with the Wittgenstein of the *Tractatus*.¹²⁴ The most straightforward version of it would be to treat universally and existentially quantified statements as, respectively, conjunctions and disjunctions of atomic predications. To say that all dogs bark is ultimately just to say that Fido barks, Sparky barks, Cerberus barks, and so on for all the dogs.¹²⁵ To say that some dog is shed hair all over the couch is to say that either Fido did, or Sparky did, or Cerberus did, and so on for all the dogs.

If we went this route, what we'd say is that the logical complexity that quantified sentences express is that of complex propositional attitudes of a conjunctive or disjunctive sort. For instance, by saying "some dog shed hair all over the couch" one would express the attitude of not being willing to deny all of the propositions that ascribe couch-shedding to various individual dogs. Let's say, for the sake of argument, that Fido, Sparky and Cerberus are the only three dogs. In Sequentese, the attitude corresponding to "Some dog shed hair all over the couch" would be expressed by the following:

⊢ Shed(fido), Shed(sparky), Shed(cerberus)

124. The idea is never expressed very succinctly in the *Tractatus*, but it can be gathered from propositions 5.5 – 5.5352. In Wittgenstein's version, quantifications aren't replaced by long conjunctions or long disjunctions but by means of a variadic (variably polyadic) truth function $N(\xi, \dots)$ which jointly denies all of the propositions it has as arguments – a kind of extendable Sheffer stroke.

125. As a matter of fact, not all dogs bark. The Basenji or African Barkless Dog can only yodel. True fact.

But this approach soon runs into some trouble. First off, it's open to serious doubt whether the mere fact that Fido, Sparky and Cerberus were the only three dogs would suffice for the above expression of Sequentese to express the existential quantification we started with. If things were different such that there was a fourth dog, then the expression would not express that some dog shed hair all over the couch, but rather than one of a more restricted set of dogs did so. We would presumably want the equivalence of "Some dog shed all over the couch" and "Fido or Sparky or Cerberus shed hair all over the couch" not to be quite as contingent as this. If there had been a fourth dog, "Some dog shed all over the couch" should have come out true if this fourth dog shed all over the couch.¹²⁶

Secondly, one might worry that we often quantify over infinite domains ("every natural number has a successor") and that on this view of quantification, a sentence like that would express a propositional attitude of infinite complexity. It is very indeed doubtful that you and I are capable of having infinitely complex propositional attitudes. Since we are clearly able to quantify over infinite domains, one might think that the propositional attitude that such a quantification consists in had better not be infinitely complex.

We shouldn't, though, overestimate the weight we should accord this second worry. One might think that, although we, human beings, couldn't get much use out of a way of quantifying that required infinitely complex propositional attitudes, such propositional attitudes are still metaphysically possible, and some possible being might have them. That may be enough to make the point that quantifying doesn't essentially require any capacities or concepts that someone who could use the logical connectives didn't already have. And if we've shown previously that the logical connectives don't commit us to logico-metaphysical structure, then we've also shown that quantificational thought doesn't. Similar points (though not about this particular topic) have been made by Joseph Melia (see for instance his 1995).

126. This particular problem does not arise for Wittgenstein in the *Tractatus* because he takes the total number of objects to be fixed as a matter of necessity. In a philosophically perspicuous language, whether or not a long conjunction expresses a fully general universal quantification could be 'read off': one could simply check whether all the names in the language were involved in the conjunction, and if they were, then all the objects would be accounted for. We're not making any of these bold assumptions, so the problem would arise for us.

I think, nevertheless, that the first worry is enough to make us cast around for some other approach. Thankfully, there is another route open, the one on which we treat the quantificational vocabulary as bringing something genuinely novel to the language.

I.II. Ramsey-Mill Quantification

For a second, more promising line of attack, I'm going to take inspiration from a suggestion Frank Ramsey made in his "General Propositions and Causality", and which he in turn attributes to Mill. This is what Ramsey says of the universal proposition:

"It always goes beyond what we know or want; cf. Mill on 'All men are mortal and 'The Duke of Wellington is mortal'. It expresses an inference we are at any time prepared to make, not a belief of the primary sort. A belief of the primary sort is a map of neighbouring space by which we steer. It remains such a map no matter how we complicate it or fill in details. But if we professedly extend it to infinity, it is no longer a map; we cannot take it in or steer by it. Our journey is over before we need its remoter parts."¹²⁷

Later on, he adds that general beliefs are "not judgments but rules for judging 'if I meet a ϕ I shall regard it as a ψ '"¹²⁸. He goes on to make various other claims about them, and among other things, he is very quick to identify universal generalisations with causal laws.¹²⁹ But what I'll do is to just take the basic Ramseyan idea – that a universal statement expresses that one is prepared to make a certain sort of judgment 'at any time' – and adapt it to my our purposes.

A universally quantified sentence, I propose, expresses a general willingness to deploy a certain simple or complex predicate: a willingness to apply it to any object

127. Ramsey 1978, p. 134.

128. Ibid., p. 137.

129. Ibid.

one is faced with or otherwise has cause to consider. Quantifiers, on this proposal, do bring something new to the table that wasn't already implicit in the connectives.

Below I'll spell out the proposal properly. But in order to do so we must first step back from the quantifiers and talk about something more basic first: the notion of predication.

I.II.I. Predication

Up to this point, logical expressivism has been a theory just of propositional logic. The basic attitudes of acceptance and rejection, as I've described them, take simple or atomic propositions as their contents. So far we've treated these as unstructured simples, and symbolised them with single letters P, Q, R, etc. But if we're going to extend expressivism to quantifiers, and we're not using the stratagem of defining away the quantifiers in the Tractarian manner, some sub-propositional structure needs to be added in. In the context of logical expressivism, there could be more than one way to go about adding in that structure. One option would be to simply introduce structure into the propositions that acceptances and rejections take as their objects – to say that acceptance and rejection are not, as previously stated, attitudes to simple things, lacking all logical structure, but rather attitudes to complexes of objects and properties. Another option, subtly but importantly different, would be to replace the attitudes of acceptance and rejection with other, more complex attitudes which take as their objects not propositions built up somehow out of propositions and objects, but the propositions and objects themselves. I'm going to take the latter approach, because it seems neater: we can avoid the complexities of having to explain what an atomic predicational proposition is, and in fact we can now altogether remove propositions from the logical expressivist story.

So how does it go? Let's first introduce two new notions: the notion of an *object* and the notion of a *concept*. Objects are the things we talk about: tables, chairs, people, molecules, planets, countries and what have you. It is a true catch-all

notion: anything whatsoever we can talk about can be treated as an object – apart perhaps from a few things that, if they were to be treated as objects, would entangle us in paradox.¹³⁰ Concepts are what we sort these objects under, when we describe them as being such-and-such or so-and-so. If one believes that Norway is a wealthy country, then one sorts the object Norway under the concept of being wealthy. On a linguistic level, one expresses this by combining, in some particular manner dictated by the syntax of one's language, a name for Norway with a predicate expressing the concept of being wealthy. I'll refer to this mental attitude of sorting an object under a concept as an *application* of a concept to an object.

Concepts, like the predicates that express them, can be single-place (monadic) or many-place (polyadic). We may take concepts to correlate with *properties*, so that the correctness of an application of a concept to an object can generally be explained metaphysically by the exemplification by that object of the property correlated with the concept. But we needn't go so far as to commit to the claim that every concept correlates with a property: perhaps in some cases the correctness of an application can be explained in some more roundabout way, in terms of properties correlated with other predicates (or perhaps in terms of properties correlated with no term of our language at all). However, the detailed metaphysics of properties and their relations to concepts and predicates is not going to be very important in what follows.

How is the act or attitude of application to be connected to the attitudes of acceptance and rejection? One straightforward option for doing that would be to simply identify the attitude of accepting an atomic (monadic or polyadic) proposition with the application of a (monadic or polyadic) concept to an object. We could then perhaps go on to identify the attitude of rejection with the application of the *complement-concept* of some concept to an object, where a concept's complement-concept is the thing under which all and only those objects are correctly sorted which are not correctly sorted under the concept in question.

130. And according to one philosophical tradition, we've got no business talking about such things anyway.

This isn't the way to go. One reason is that the notion of a complement-concept is somewhat awkward and unappealing, being defined as it is in terms of negation. I'm not sure it's illegitimate, but I'd rather avoid it. But another, more important reason is that this straightforward identification of acceptance and application would ultimately make our account of quantification more complicated and ugly than it needs to be. I won't explain that here - it will become clear below how and why that is the case.

Instead I'm going to draw a more indirect connection between acceptance and application, and between rejection and application. I'm going to understand the acceptance of an atomic proposition as a type of pro-attitude to an application, a sort of ruling-in of an application. If one accepts an atomic proposition, then with regard to some concepts and some object(s), one regards it as appropriate to apply the concept to the object. That one regards it as appropriate to apply a concept to an object does not imply that one does so apply it – the two are distinct attitudes. It might be thought, of course, that there is something weird in doing the former but not the latter, or vice versa. To apply a concept to a concept to an object and not accept the corresponding atomic predication would indeed be an outright failure of rationality. But the converse needn't always be. I think there are cases in which one regards an application as appropriate and doesn't make the application, with perfect legitimacy. I'll say more about that in a bit, below.

Now of course one might expect that in the normal run of things the acceptance of an atomic proposition does bring along with it the actual application of the concept in question to the object(s) in question, and vice versa. The two might be psychologically indistinguishable in a lot of cases. Nevertheless, it is important that there be the conceptual possibility of them coming apart: it's going to be crucial in the case of quantification.

What about rejection? Rejection, as one might expect, will be a con-attitude equal and opposed to acceptance: if one rejects a certain atomic proposition, then one refuses, or rules out, the application of the concept involved to the object(s) involved. That one rules out a certain application does not imply that one doesn't actually make that particular application – but in doing so, one would be doing

something irrational. In general though, one might expect that when a certain application is ruled out, the application doesn't occur. The converse should not be expected to hold generally: there are always going to be lots of cases in which a subject doesn't apply a specific concept to a particular object simply because they haven't thought about those concepts or those objects, and might never have cause to. In those cases, one wouldn't expect the subject to rule out the applications in question.

Having said something about what I take the attitudes relating to predication to be, we can now say something more specific about quantification.

I.II.II. Quantification

Syntactically, quantifiers are operators that bind variables. If we take away a quantifier from a sentence, what we get is an 'open' sentence with unbound variables in it. To be able to explain what attitudes are expressed by quantified sentence, it will be useful to first say something about what an open sentence, i.e. one with the quantifier not yet attached, can be taken to express.

If we take a sentence of not-yet-quantified predicate logic, i.e. one that either is an atomic (monadic or polyadic) predication or built up out of atomic predications by means of logical connectives, and if we replace one of the names in that sentence with a variables throughout, what we get is an open sentence. This open sentence, unlike the sentence from which it is derived, doesn't make a descriptive claim that we can assess for correctness. By replacing the name with an unbound variable, we've produced a pseudo-sentence that isn't about anything in particular. But it still expresses something.

Take, for example, the complex sentence "If Katie is a philosopher, then Katie has a beard". We replace both instances of "Katie" with a variable "x" and we get "If x is a philosopher then x has a beard". That new sentence isn't one we can judge to be true or false *simpliciter*, given that we don't know what the sentence is about; we don't know what x is. But we can judge this sentence to be true or false *of things*. If

the original sentence was true, then Katie is one of the things that this open sentence is true of. It may also be true of Jim, false of Mary, etc.

In the logical expressivist framework, open sentences are taken to express something quite like a pattern of attitudes – something that can be ruled in or ruled out. But it's not exactly a pattern of attitudes because it's not fully determinate. Take the above example again. The logical expressivist takes the sentence “If Katie is a philosopher, then Katie has a beard” to express the ruling-out of a particular pattern: the pattern of accepting the application of being a philosopher to Katie and rejecting the application of having a beard to Katie. Let's take that pattern and abstract away from Katie: we get the 'pattern' of accepting of some thing that it is a philosopher and rejecting of that same thing that it has a beard. Now just as we couldn't judge of an open sentence whether it was true or false, we can't judge this pattern to be one that one ought to rule in or rule out. But we can take some object and judge whether, if the attitudes in the pattern were about that object, they would be appropriate. So if Jim is a bearded philosopher, then if the attitudes in the pattern were attitudes towards Jim, we should rule out that pattern. If Katie is a beardless philosopher, and if the pattern is about her, we should rule it in.

I'm going to call the sort of pseudo-pattern that you get by taking a pattern and abstracting away from some object that figures in the attitudes a *condition*. A condition is something which, I will say, objects can meet or fail to meet. Jim fails to meet the condition above; therefore, if it were turned into a pattern of attitudes about him, it ought to be ruled out. That means that the sentence “If Jim is a philosopher, Jim has a beard”, which rules out this very pattern, expresses a correct attitude; we could call it true for that reason¹³¹.

Patterns can be ruled in or ruled out. Something similar can happen with conditions, or so I'll propose. Ruling out a condition C is an attitude which involves ruling out the patterns which would result from inserting any object whatsoever in the 'open' parts of C. Ruling out the pattern of (accepting that Katie is a philosopher and rejecting that she has a beard) is to have an opinion on Katie in particular.

131. For some more detailed discussion of the notions of correctness and truth, see the conclusion; I take them as fairly intuitive at this point.

Ruling out the corresponding condition is to have an opinion on objects generally: take any one of them, and one would rule out accepting that it's a philosopher and rejecting that it has a beard. Similarly, ruling in a condition involves ruling in the patterns that would result from inserting any object whatsoever into the condition.

I want to understand the kind of generality that's involved in ruling out and ruling in conditions in the way suggested by Ramsey. I do not take it that, generally, the way one rules out/rules in a condition is by first going through all the objects and ruling out/ruling in the corresponding patterns for each and every one of them. That may be a way way of ruling out/ruling in a condition, but not a very common way, I would think. Rather, I'm proposing that, just as one can have an opinion on the object that is Katie, one can have an *indifferent* attitude to objects in general, which consists in a disposition to make certain specific judgements about any object whatsoever that might cross one's path or one's thoughts. To believe that everything is physical is to be willing to judge, of any object that comes to one's attention, that it is physical. The kind of ruling in and ruling out that one can do to conditions I take to involve such an indifference to the identities of objects.

It is these attitudes of ruling in and ruling out conditions (as opposed to patterns) that I take to be the attitudes expressed by quantified sentences. As we'll see in a little bit when we bring out the Sequentese formalism again, I will take sequents to sometimes express (when marked in a certain way) the ruling out of a condition rather than of a pattern. Those sequents correspond to sentences of quantified predicate logic. Attitudes of ruling in and ruling out conditions are going to be treated just the same way as attitudes of ruling in and ruling out patterns were: they have the same syntactic role, they embed in the same way, and they obey the same inference rules.

I.III. Correctness

The expressivist story has now become a little more complicated, in at least two ways. First, there are now new attitudes in the picture, and also new objects for

these attitudes: conditions. Second, the simple propositions that were still left over at the end of chapter III have now dropped out of the picture, to be replaced with objects and concepts, and attitudes involving these.

The metasemantic story outlined in chapter IV is supposed to extend smoothly to the case of the quantifiers: there should be nothing in that story that is particular to the case of the connectives. But in that story we made use of the notion of correctness conditions, and those were defined in a way that presupposed that there were simple propositions that were true or false. Also, the definition made no mention of the notion of a condition, as introduced in this chapter. So here is the definition of correctness, as extended to these new notions:

Correctness 0-a: An attitude of accepting the application of a concept F to a sequence of objects $\alpha, \beta, \gamma, \dots$ is correct iff that sequence falls under F ;

Correctness 0-b: An attitude of rejecting the application of a concept F to a sequence of objects $\alpha, \beta, \gamma, \dots$ is correct iff that sequence does not fall under F ;

Correctness 1-a: The ruling-in of a level-1 pattern is correct iff all of the attitudes in the pattern are correct;

Correctness 1-b: The ruling-out of a level-1 pattern is correct iff some of the attitudes in the pattern are not correct;

Correctness 1-c: The ruling-in of a level-1 condition is correct iff all of the attitudes in the pattern are correct for all objects;

Correctness 1-d: The ruling-out of a level-1 condition is correct iff some of the attitudes in the pattern are not correct for some objects;

Correctness 2-a: The ruling-in of a level-2 pattern is correct iff all of the attitudes in the pattern are correct;

Correctness 2-b: The ruling-out of a level-2 pattern is correct iff some of the attitudes in the pattern are not correct;

Correctness 2-c: The ruling-in of a level-2 condition is correct iff all of the attitudes in the pattern are correct for all objects;

Correctness 2-d: The ruling-out of a level-2 condition is correct iff some of the attitudes in the pattern are not correct for some objects;

...

...

Correctness N-a: The ruling-in of a level-n pattern is correct iff all of the attitudes in the pattern are correct;¹³²

Correctness N-b: The ruling-out of a level-n pattern is correct iff some of the attitudes in the pattern are not correct;

Correctness N-c: The ruling-in of a level-n condition is correct iff all of the attitudes in the pattern are correct for all objects;

Correctness N-d: The ruling-out of a level-n condition is correct iff some of the attitudes in the pattern are not correct for some objects;

Interestingly, the notion of correctness is now no longer reliant, definitionally, on the notion of truth. But we can still define a notion of (sentence-)truth in terms of correctness. In the previous chapter we noted that the notion of correctness could be extended to sentences: a sentence can be said to be correct iff it is used to express an attitude which is correct, and incorrect iff it is used to express an attitude which is incorrect. We can let sentence-truth function in the same way: a sentence can be said to be true iff it is used to express an attitude which is correct, and false iff it used to express an attitude which is incorrect.

I.IV. On the way to Predicate Sequentese

How do all these new attitudes fit in with the Sequentese formalism? Let's first talk through an example, before dealing with it in general and in formal detail. Take the following fairly simple sentence of Sequentese:

$P, (T \blacktriangleright U) \blacktriangleright Q, (R \blacktriangleright S)$

Now imagine that the various propositional letters represent monadic predications, as follows:

$F(a), (G(b) \blacktriangleright H(b)) \blacktriangleright G(a), (H(b) \blacktriangleright G(b))$

132. Note that by the definitions of the different levels of patterns, the attitudes of the patterns will always be of a lower level, and so their correctness conditions will have been defined at that point.

This would express the following:

“Ruling out accepting the application of F to a and (ruling out accepting the application of G to b while rejecting the application of H to b) while rejecting the application of G to a and (ruling out accepting the application of H to b while rejecting the application of G to b).”

It's convoluted, but it's understandable and it should be clear how it was arrived at. But now let's replace some of the names with variables:

$$F(x), (G(b) \blacktriangleright H(b)) \blacktriangleright G(x), (H(b) \blacktriangleright G(b))$$

Now the sequent would express a condition, and we can describe it as follows:

“Ruling out accepting the application of F to the object in question and (ruling out accepting the application of G to b while rejecting the application of H to b) while rejecting the application of G to the object in question and (ruling out accepting the application of H to b while rejecting the application of G to b).”

Again this is similarly convoluted, but it should be clear how we got there: we took out a name and replaced it with an indefinite description, rendering the whole thing indefinite in that respect.

Now we are on our way to understanding how quantification might work in Sequentese. For if we had a way of *universally closing* an open sentence like the above, the result could function as the expression of the ruling-out of the condition expressed by the corresponding open sentence. As a formal device indicating universal closure, I'm going to subscript pointers with a variable. This should be taken to indicate that all of the variables corresponding to that subscript that appear within the scope of that pointer are universally closed. For example:

$$F(x), (G(b) \blacktriangleright H(b)) \blacktriangleright_x G(x), (H(b) \blacktriangleright G(b))$$

This would then express the following attitude:

“Ruling out, *for any x*, accepting the application of F to x and (ruling out accepting the application of G to b while rejecting the application of H

to b) while rejecting the application of G to x and (ruling out accepting the application of H to b while rejecting the application of G to b)."

And if we replace the "b" with another variable, we can bind that too by subscripting the same pointer:

$$F(x), (G(y) \blacktriangleright H(y)) \blacktriangleright_{x,y} G(x), (H(y) \blacktriangleright G(y))$$

This will express:

"Ruling out, *for any x and for any y*, accepting the application of F to x and (ruling out accepting the application of G to y while rejecting the application of H to y) while rejecting the application of G to y and (ruling out accepting the application of H to y while rejecting the application of G to b)."

That should give an idea of how things are going to go. But now, let's state properly how Predicate Sequentese works.

II. Sequentese with Predicates: Syntax

II.I. Predication

To extend Sequentese to capture the behaviour of quantifiers, we must first allow it 'see' some structure below the propositional level: subject-predicate structure.

So far Sequentese has represented the contents of propositional attitudes simply with proposition-letters: "P", "Q", "R", etc. When "P" crops up on the left-hand-side of a pointer, then it is, so to say, in 'acceptance position', and if "Q" crops up on the right, then that letter is in 'rejection position'. This is to say that when we translate "P \blacktriangleright Q" into English, we write "the attitude of ruling out to *accept* P while *rejecting* Q. The position of a propositional letter in a sequent dictates the kind of simple propositional attitude (acceptance or rejection) that it comes to represent.

Let's now introduce some new vocabulary. Let's have some *predicate letters*, "F", "G", "H", etc. We can generate as many as we like using numbered subscripts. Let's also introduce some *object variables* "x", "y", "z", "x₁", "x₂", "x₃". And let's say we can combine these, using the vertical stroke "|", to create expressions like "F|x" and "G|y". These compound expressions – call them *predicate ascriptions* or just *ascriptions* for short – will behave exactly like sentence letters, syntactically: they can appear on the sides of a pointer, separated from other ascriptions or propositional letters by commas. There can be unary ones, like "F|x", binary ones, written "R|x|y" and indeed n-ary ones for any n.

An n-ary predicate ascription will represent the attitude of applying an n-ary predicate to an object or objects. So "F|x" will represent the act or attitude of applying the predicate that "F" stands for to whichever object is assigned to the variable "x". (As a bit of meta-logical terminology, let's use "g(x)" to abbreviate "the object assigned to "x"".)

A predicate ascription, like a propositional letter, can appear on the left or on the right of a pointer, that is, on the acceptance side or the rejection side. When it appears on the left-hand-side, it will represent the *acceptance* of the application it represents, and when it appears on the right-hand-side, it will represent the *rejection* of the application it represents. So we would translate "F|x ► G|y" as follows:

"Ruling out accepting the application of F to x while rejecting the application of G to y".

For brevity's sake, we'll sometimes just opt for half-way translations, i.e: "the attitude of being unwilling to accept F|x while rejecting G|y", on the understanding that such half-way translations are to be expanded in the manner just above.

II.II. Open and Closed Sequents

When we have a sequent like $F|x \triangleright P$ we cannot tell whether the attitude it expresses would be an appropriate one to have unless we know which object is assigned to the variable x . Until we have such an assignment, we don't really know which attitude is being expressed. So let's call a sequent like $F|x \triangleright P$, i.e. one which leaves the value of the variable undetermined, an 'open' sequent: more particularly, this sequent is one that's 'open in x '.

So how do we get a closed sequent? Let's allow our pointers to be subscripted with variables, like this: \triangleright_x . When a pointer is subscripted in this way, we can regard it as expressing the universal closure its open counterpart (i.e. the same sequent without the subscript). So if $F|x \triangleright P$ translates as:

"Ruling out accepting $F|x$ while rejecting P "

Then $F|x \triangleright_x P$ translates as:

"Ruling out, for all (assignments of objects to) x , accepting $F|x$ while rejecting P ".

To bind an open sentence which is open in more than one place, one would need multiple subscripts. Sequentese allows pointers to be multiply subscripted. So to bind $R|x|y \triangleright P$, one could write $R|x|y \triangleright_{x,y} P$. That would translate as:

"Ruling out, for all x and y , accepting $R|x|y$ while rejecting P ."

One might worry that in cases where a pointer is multiply subscripted, it is not clear which of the 'bindings' takes scope over the other. This is not a problem. If Sequentese had multiple styles of binding variables, one say for universally closing a sequent and one for existentially closing it, the order would matter a great deal. But note that in predicate logic, a sentence with one universal quantifier taking outer scope and another taking next-to-outer scope is equivalent to the sentence with those two quantifiers switched around. In other words, the order of two universal quantifiers (and similarly two existential quantifiers) is insignificant. For closely related reasons (and this will become clearer below) there is no need to indicate an order of precedence between bindings in Sequentese.

So what is a well-formed sequent of Predicate Sequentese? Relying on the definition of a well-formed sequent of Propositional Sequentese, we can specify this as follows:

(i) Any sequent of Propositional Sequentese is a well-formed formula of Predicate Sequentese.

(ii) Any sequent which is the result of replacing, in a well-formed formula of Propositional Sequentese, any sentence-letter (P, Q, R, etc.) with an ascription ($F|x$, $G|y$, $R|x|y$, etc.) is a well-formed formula of Predicate Sequentese.

(iii) Any sequent which is the result of taking a sequent of the sort described under clause (ii) above, and subscribing one of the pointers therein with one or more variables also occurring in ascriptions within its scope, is a well-formed formula of Predicate Sequentese.

(v) Nothing else is a well-formed formula of Predicate Sequentese.

III. Sequentese with Predicates: Translation Manual

How do we extend our Sequentese-PL translation manuals to include these new bits of formalism? The target will now be something different, because we won't be translating to and from the language of classical first-order propositional logic, but to and from the language of classical first-order predicate logic (without names or identity – for now). So we'll have expressions like " $F(x)$ " and " $G(y)$ " to deal with, and the quantifiers " \forall " and " \exists ".

Atomic open predications like " $F(x)$ " are obvious: we translate the predicate term by whatever its counterpart is in Sequentese (let's assume it will be a homograph) and the variable by a Sequentese one (assume, again, a homograph). Instead of the bracket notation, we have the Sequentese " $|$ ", as above. If the sentence of predicate logic we're translating happens to be an open sentence with no quantifiers, then we can simply translate the atomic predications and then proceed

with the rest of the sentence exactly as we would do by the lights of the translation manual of chapter III. The other direction will be similar. We can translate a fully open sequent by means of the old translation manual, and then replace, in our translation, expressions like “ $F|x$ ” with expressions like “ $F(x)$ ”.

But what if there are quantifiers? Let’s look at Sequentese-to-predicate-logic first. If we have a universally closed sentence like “ $F|x \multimap_x P$ ”, we do the following: ignoring the subscript at first, we translate the sentence as if it was an open one, as above. So we’d get “ $F(x) \rightarrow P$ ”. Then we enclose that in brackets and put a universal quantifier in front of it, binding the “ x ”, like so: “ $\forall x(F(x) \rightarrow P)$ ”. The same applies if the universally closed sequent is embedded, except that there’d be some more translating left to do, of the sequent that embeds it. That’s all there is to it, since there’s only one way of making a sequent closed.

The other way around, predicate-logic-to-Sequentese, is a little more involved. One case is fairly easy, that of translating sentences with outer-scope universal quantifiers. In this case we first ignore the quantifier, translate its open counterpart, and then subscript the outer-scope pointer in the resultant translation with the appropriate variable. But other cases are less obvious. What about iterated universal quantifiers, such as in the sentence “ $\forall x \forall y (F(x) \rightarrow G(y))$ ”? In case of iterated universal quantifiers, we first take away the quantifiers first, and translate the resultant open sentence. Then we reintroduce both quantifiers to our translated open sentence in one fell swoop by subscripting the pointer in “ $F|x \multimap G|y$ ” *twice*, like so: “ $F|x \multimap_{x,y} G|y$ ”. For the sake of having a determinate translation manual, I’ll stipulate that the pointer is to be subscripted in the order of the original quantifiers. But really, it doesn’t matter.

How about existential quantifiers? These will have to be translated by a step-wise procedure. Let’s say we have a sentence like “ $\exists x(F(x))$ ”. Given the duality of the universal and existential quantifiers, this is equivalent to $\neg \forall \neg (F(x))$. This is something we can translate: building up the translation in a step-wise manner, according to our manual, we’d get this:

“ $F|x$ ”

"F|x ▶"

"F|x ▶_x"

"(F|x) ▶_x ▶"

Finally, what about embedded quantifiers? How about the sentence " $\exists x(Fx \wedge \forall y(Gx \rightarrow Rxy))$ "? The recipe is: start at the smallest translatable units (predications or single sentence-letters) translate those, and then add in the logical structure step-wise, adding in subscripts to pointers at the earliest point we can do so without leaving any variables unbound. So in the above, we first translate "Fx", "Gx" and "Rxy". They will become, respectively, "F|x", "G|x" and "R|x|y". The smallest bit of logical structure we can then implement is the conditional: this gives us "G|x ▶ R|x|y". Then we're allowed to subscript that pointer with a "y", for there are no other "y"s in the sentence which would thereby remain unbound. That gets us "G|x ▶_y R|x|y". Then we're ready to translate the conjunction. Conjunctions, as we've seen in Chapter III, are a bit unwieldy, but this is what we get (in effect we're exploiting the equivalence between " $A \wedge B$ " and " $\neg(\neg A \vee \neg B)$ ":

(▶ ((G|x ▶_y R|x|y) ▶), (F|x ▶)) ▶

Then we just have to add the existential quantifier to our translation. Using our workaround above, we first 'negate' the sequent:

((▶ ((G|x ▶_y R|x|y) ▶), (F|x ▶)) ▶) ▶

We then subscript the outer-scope pointer like so:

((▶ ((G|x ▶_y R|x|y) ▶), (F|x ▶)) ▶) ▶_x

And then we 'negate' the result again:

((▶ ((▶ ((G|x ▶_y R|x|y) ▶), (F|x ▶)) ▶) ▶_x) ▶

These recipes give us, in a nutshell, all we need to translate everything we might want to translate.

IV.I. Sequentese with Predicates: Inference Rules

Translation is one thing, but in chapter 3 we also introduced a proof system for the Sequentese formalism. We can do the same for the quantified version of Sequentese, by giving some new inference rules to govern the new symbols. After motivating these rules, we can then show that they end up giving us the classical predicate logic that we know and possibly love.

First, however, we must say something about what proofs look like in Predicate Sequentese. One odd feature of Predicate Sequentese is that it is a language with no names. It has only variables, which can appear in open and closed sequents (both of which are regarded as well-formed).

Also, in Predicate Sequentese, open sequents are regarded as well-formed. They can also figure in arguments. A Sequentese proof is a series of sequents, open or closed, the last one of which is the conclusion. A valid proof of Sequentese is one in which every line is either a premise or derived from one or more premises by an inference rule of Sequentese (Propositional or Predicate). Any line of a Sequentese proof may be an open sequent, but the conclusion of a proof may only be an open sequent if one of the premises is.

These are oddities of Predicate Sequentese. But they are complementary oddities. Predicate Sequentese's lack of names needn't be a hindrance – we can still do all the reasoning we want Sequentese that we want to do. If I want to reason from the premise that Lucky is a dog, what I do is take the open sequent $\text{IsADog} \mid x$, and assign Lucky to the x . Then, in reasoning from this sequent, I will be reasoning about Lucky. When a Sequentese argument has an open premise, then the open variables in that premise will be treated as, effectively, names. We'll read the argument under the assumption that some object has been assigned to that variable and that this object satisfies the premise(s). Wherever else that variable occurs in the argument, unbound, it is read under that same assignment.¹³³

133. Another option worth mentioning, if one did want to avoid having open variables in premises, is to replace them with Russellian descriptions: instead of just assigning the x to Lucky, I could have a quantified premise which manages to talk about Lucky in virtue of him being the only object that satisfies the description in the sentence. E.g. I could say that there is a thing which is my pet, and that anything that is my pet is identical to it, and that it's a dog. I think, however, that the above approach is elegant enough.

Of course, in lots of cases where we reason in formal systems, we aren't actually concerned with identified objects. The important thing is that we treat the variables open in premises as having assignments, and as having those same assignments throughout. That this is the important thing is even clearer when a variable occurs unbound in a Sequentese argument but it does *not* occur in the premises. This may happen, for as we'll see, there is a rule of Predicates Sequentese which creates an open sequent from a closed one. As long as those newly unbound variables remain unbound, we treat them as having a constant assignment. But it matters not at all in this case what the assignment is – indeed, the less we know about it the better.

Keeping in mind that this is how open sequents are to be read in Sequentese proofs, we can present some inference rules. I'll first introduce, and tinker with, some likely-looking inference rules. Then I'll say something more about how I take these to be motivated.

There's basically only one symbol that we need to legislate for, the subscripted pointer. I suggest, first, the following rule:

- i) $\phi \blacktriangleright_{\alpha} \psi$
- OPEN
- j) $\phi \blacktriangleright \psi$

This 'opening rule' is a bit like universal instantiation: if ϕ or ψ have the variable α occurring in them, the resulting sequent will be an open one. But assuming that the original closed sequent is in good order, the resultant open sequent will be in good order on any assignment of an object to the open variable. This rule looks pretty good, then.¹³⁴

Another rule:

- i) $\phi \blacktriangleright \psi$
- CLOSE
- j) $\phi \blacktriangleright_{\alpha} \psi$

134. I'm using Greek lowercase letters here to stand in indifferently for atomic proposition-letters, atomic predications and embedded sequents.

This is a 'closing rule' that does the reverse of the rule above. But this rule must be restricted if it is to be acceptable. If either ϕ is open in α , or ψ is, or both are, then α would need to be, so to say, an 'arbitrary' variable, i.e. not one about which we had special information – otherwise the generalisation wouldn't be legitimate. But how do we pin that down? It's fairly straightforward: we require that α is not open in any premise-line.

The reason for this is that, as noted above, a proof in predicate Sequentese, if it has open variables in the premises, does not signify anything if the variables don't receive an assignment. So when a Sequentese proof would actually be used by someone as a guide to reasoning, it will be under an assignment of objects to those variables (see section VI for more on this). If one assigns a specific object, say the moon, to a variable β , then β will for the purposes of that argument be a name for the moon, when it occurs unbound. So the premises, under this assignment, will contain specific information about the moon. If I have derived something which is, in this sense, specifically about the moon, I will obviously not be allowed to base completely general conclusions about all objects on this. If, however, β was not open in the premises, then it was introduced by the OPEN rule (this will be the only rule that introduces unbound variables). It will thus be assigned an arbitrary, unidentified object, i.e. an object about which we know nothing that sets it apart from other objects, and will behave like a name for that object during the time in which it is unbound.

These two are the most important rules we need to govern the pointer subscripts. They are pretty straightforward analogues of the kinds of introduction and elimination rules one gets for the universal quantifier in Natural Deduction and Sequent Calculus settings, if one imagines that, in the case of OPEN, we replace the opened variables with constants, and if one reads, in the case of CLOSE, 'constants in the premises' for 'open variables in the premises'. I take it not to stand in need of formal proof that these rules are classical.

We also need to say something about rules introduced in chapter III: how do these behave when some pointers in the sequents we're manipulating are subscripted?

There is some trickiness here. Take for instance the REM(r) rule:

i) $\Theta, A \triangleright \Lambda$

----- REM(r)

j) $\Theta \triangleright (A \triangleright B), \Lambda$

In the sequent above the line, there's one pointer. In the one below the line, there are two. Suppose the sequent above the line is universally closed in α : " $\Theta, A \triangleright_{\alpha} \Lambda$ ". When we apply REM(r), where does the subscript go? To the inner-scope pointer? To the outer-scope one? It seems it can't always go to the inner-scope button, because there may be some ascriptions involving α that would then be left out of its scope, becoming open as an unintended consequence. It can go with the outer-scope pointer without such problems arising, but as it turns out this makes the rule a lot less powerful.

To illustrate, take the following argument:

(i) $A, F|x \triangleright_x G|x$

(ii) $A \triangleright_x (F|x \triangleright D), G|x$ REM(r)

Line (i) rules out, for all x , accepting A , accepting $F|x$, and rejecting $G|x$. Line (ii) rules out, for all x , accepting A and ruling in (accepting $F|x$ while rejecting D) and rejecting $G|x$. If the x had gone with the smaller-scope pointer in (ii), it would have ruled out accepting A , ruling in, (for all x , accepting $F|x$ and rejecting D), and rejecting $G|x$. It wouldn't be clear, in that case, what the value of the x in " $G|x$ " was – it wouldn't have been the same x that the "for all x " clause mentioned. So that would have been bad.

But contrast the following argument:

(i) $A, F|x \triangleright_x B$

(ii') $A \triangleright_x (F|x \triangleright D), B$ REM(r)

Line (i) rules out, for all x , accepting A , accepting $F|x$ and rejecting B . Line (ii') rules out, for all x , accepting A and ruling in (accepting $F|x$ while rejecting D) and rejecting B . That's fine – if one had the former attitude, the latter would have been

appropriate. But suppose the x had gone with the smaller-scope pointer in (ii'), like so:

$$(i) \quad A, F|x \triangleright_x B$$

$$(ii'') \quad A \triangleright (F|x \triangleright_x D), B \quad \text{REM}(r)$$

Line (ii'') rules out accepting A , ruling in, for all x , accepting $F|x$ and rejecting D , and rejecting B . Unlike in the former case, there's no trouble with unbound variables popping up. (ii'') is *also* something that follows from (I), though distinct from it. And of course, I might well be in the position where I want to derive (ii''') instead of (ii'). But it's unclear, if we take this 'safer' version of $\text{REM}(r)$, how I'm going to get it.

Here's a solution: we can opt for the safer, less powerful version of $\text{REM}(r)$, and introduce an additional rule in predicate Sequentese that makes up for the weakness of the rule in those cases where a stronger rule wouldn't steer us wrong. Here's that rule:

$$i) \quad \phi \triangleright_{\alpha_1} (\chi \triangleright \psi)$$

----- SCOPE (in)

$$j) \quad \phi \triangleright (\chi \triangleright_{\alpha_1} \psi)$$

This rule comes with restrictions: it allows us to move a subscript from an outer-scope pointer to an inner-scope pointer only if by doing so, no variables that were bound thereby become unbound. The subscript can go narrow-scope as long as no unboundedness results. Whenever we apply the Sequentese inference rules that multiply pointers, to wit LEM , $\text{REM}(r)$ and $\text{REM}(a)$, we might get situations like the above, where a new pointer comes into the sequent. In those cases SCOPE may be useful.

SCOPE will also have a converse rule, which allows one to move a subscript from an inner-scope pointer to an outer-scope one:

$$i) \quad \phi \triangleright (\chi \triangleright_{\alpha} \psi)$$

----- SCOPE (out)

$$j) \quad \phi \triangleright_{\alpha} (\chi \triangleright \psi)$$

This rule, and also the other rules of propositional Sequentese which are such that they make some pointers disappear, i.e. REX, LEX(a) and LEX(r), lead to the possibility of variables that are apparently doubly bound. With REX, LEX(a) and LEX(r), if any of the pointers in the above-the-line sequent are subscripted, the subscript will have to attach to the broader-scope pointer that's left over after the rule has been applied. There are some situations in which this threatens to introduce doubly bound variables, though, as in the following:

$$\text{i) } (\blacktriangleright_{\alpha} \phi) \blacktriangleright (\blacktriangleright_{\alpha} \psi)$$

----- REX

$$\text{j) } (\blacktriangleright_{\alpha} \phi) \blacktriangleright_{\alpha} \psi$$

In line (i), the variable α crops up on both sides of the outer-scope sequent, and harmlessly, because neither of the pointers that they are subscripted under are within each other's scope. Once we apply REX, though, any α that occurs in ϕ would appear to be bound by the subscripts of both pointers. Similar effects can occur with the LEX rules, and also with applications of the SCOPE (out) rule. What should we do about situations like these?

We don't have to do anything about it, as it turns out: we can simply stipulate that the convention is that when a variable is ostensibly doubly bound, it is in actual fact bound only by whichever pointer has the narrowest scope. This will take adequate care of all the above situations.

SCOPE (in) and SCOPE (out) are more of syntactic than semantic significance. If we translate the premise- and conclusion-schemata into predicate logic, they encode transitions between logically equivalent sentences quantified sentences, like the following (and its converse):

$$\text{i) } P \rightarrow \forall x F(x)$$

$$\text{j) } \forall x (P \rightarrow Fx)$$

I take it, again, that it does not stand in need of formal proof that these rules do not compromise the classical character of Predicate Sequentese.¹³⁵

IV.II. Justifying the Inference Rules.

The new inference rules of Predicate Sequentese are of a somewhat different sort than those of Propositional Sequentese. Syntactically speaking, all of these rules govern the placement of subscripts in sequents. None of them take us from a sequent of a certain form to a sequent of a strikingly different form. None of them make pointers appear or disappear. In fact, OPEN and CLOSE are the only rules were which are of more than mere syntactic significance. They are justifiable in terms of the attitudes of open and closed sequents express. As for OPEN: if one rules out some condition, then it seems that ruling out a pattern corresponding to that condition is indeed rationally allowed, and indeed that the ruling-in of that pattern is incoherent.

CLOSED is perhaps a little trickier. It may not seem quite as blindingly obvious that if one rules out a pattern involving an arbitrary object, this licenses the ruling-out of the corresponding condition. In the abstract, one might think that by proving something for a single object, arbitrary or not, one is not proving it for every object.

135. The lack of names in Sequentese has a few odd effects, such as that exchanging bound variables for other bound variables (i.e. the equivalent of going from $\forall xF(x)$ to $\forall yF(y)$ in predicate logic) is rather difficult, given that OPEN (the Sequentese equivalent of Universal Instantiation) does not remove the variable. In fact, I do not have an argument to the effect that it will always be possible. But I do not take this to be a problem with respect to the aimed-for equivalence between Predicate Sequentese and classical predicate logic. For one thing, the Sequentese argument that takes one from $T(\forall xF(x))$ to $T(\forall yF(y))$ or vice versa could just be taken to be a trivial one-line argument: one translates " $\forall xF(x)$ " to Sequentese, and then immediately translates that sequent back as " $\forall yF(y)$ ". I have not encountered any situations where one might want to exchange variables while reasoning within Sequentese, but if it were necessary then there would always be the option of introducing an extra inference rule:

$$\begin{array}{l} \text{i) } \phi(\alpha) \blacktriangleright \alpha \psi(\alpha) \\ \text{-----} \\ \text{j) } \phi(\beta/\alpha) \blacktriangleright \beta \psi(\beta/\alpha) \end{array} \quad \text{SWITCH}$$

SWITCH would be a harmless addition to the system.

But I think such qualms would reflect an insufficient appreciation of arbitrariness. An arbitrary object is not just a random object. Arbitrariness is ultimately an epistemic notion: an arbitrary object is one about which we know nothing that sets it apart from other objects. Consequently, anything we manage to prove about an arbitrary object wasn't proved on the basis of anything which sets that object apart from other objects. Consequently, whatever we prove about such an object, we prove on the basis of features which that object does have in common with other objects. And so what holds of an arbitrary object holds generally.

V. Sequentese with Predicates is Standard Predicate Logic

The classical predicate calculus (as it's described in Sider (2010)) introduces one new inference rule and two new axioms, in addition to the rule and axioms of the propositional calculus used in chapter III. The inference rule is Universal Generalisation:

$$\begin{array}{l} \phi \\ \text{----- UG} \\ \forall\alpha(\phi) \end{array}$$

The two axioms are:

$$\begin{array}{l} \text{(PC1): } \forall\alpha(\phi) \rightarrow \phi(\beta/\alpha) \\ \text{(PC2): } \forall\alpha(\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \forall\alpha(\psi)) \end{array}$$

The “ (β/α) ” bit in PC1 represents the replacement of the variable α by a term β throughout. In the case of PC2, a restriction applies: ϕ cannot be open in α . And while Sider sets the issue aside, it seems that a workable version of the predicate calculus would impose some restriction on applications of UG. One such restriction could be that the term (if any) in ϕ that gets replaced by a variable in $\forall\alpha(\phi)$ cannot occur in any premises that ϕ was derived from – in other words, it needs to be an arbitrary term. I'll interpret UG as so restricted.

Can we show that this rule and these axioms hold in predicate Sequentese? We can. Take UG:

$$T(\phi) = \blacktriangleright \phi$$

$$T(\forall\alpha(\phi)) = \blacktriangleright_{\alpha} \phi$$

The latter sequent follows from the former by one application of CLOSE, provided α is not open in any premises that ϕ is derived from; but this restriction will hold, if α in " $\forall\alpha(\phi)$ " is not replacing any term that's in a premise that ϕ is derived from. How about axiom PC1?

$$T(\forall\alpha(\phi) \rightarrow \phi(\beta/\alpha)) = (\blacktriangleright_{\alpha_1} \phi) \blacktriangleright (\blacktriangleright \phi)$$

- $(\blacktriangleright_{\alpha_1} \phi) \blacktriangleright (\blacktriangleright_{\alpha_1} \phi)$ Reflexivity
- $(\blacktriangleright_{\alpha_1} \phi) \blacktriangleright_{\alpha_1} \phi$ REX

Note that ϕ is ostensibly doubly bound, but in fact only bound by the leftmost pointer.

- $(\blacktriangleright_{\alpha_1} \phi) \blacktriangleright \phi$ OPEN
- $(\blacktriangleright_{\alpha_1} \phi) \blacktriangleright (\blacktriangleright \phi)$ REM(a)

□

And PC2:

$$T(\forall\alpha(\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \forall\alpha(\psi))) = (\phi \blacktriangleright_{\alpha_1} \psi) \blacktriangleright (\phi \blacktriangleright (\blacktriangleright_{\alpha_1} \psi))$$

$$8) (\phi \blacktriangleright_{\alpha_1} \psi) \blacktriangleright (\phi \blacktriangleright_{\alpha_1} \psi) \quad \text{Reflexivity}$$

In this case we pick a ϕ such that no part of it is open in α , as in the PC2 axiom.

$$9) (\phi \blacktriangleright_{\alpha_1} \psi), \phi \blacktriangleright_{\alpha_1} \psi \quad \text{REX}$$

$$10) (\phi \blacktriangleright_{\alpha_1} \psi), \phi \blacktriangleright_{\alpha_1} (\blacktriangleright \psi) \quad \text{REM(a)}$$

$$11) (\phi \blacktriangleright_{\alpha_1} \psi), \phi \blacktriangleright (\blacktriangleright_{\alpha_1} \psi) \quad \text{SCOPE (in)}$$

Note that we can do this only because we know that ϕ is not left open in α by the application of SCOPE (in).

$$12) (\phi \blacktriangleright_{\alpha_1} \psi), \phi \blacktriangleright (\phi \blacktriangleright (\blacktriangleright_{\alpha_1} \psi)) \quad \text{REM(a)}$$

13) $(\phi \triangleright_{\alpha_1} \psi) \triangleright (\varphi \triangleright (\triangleright_{\alpha_1} \psi)), (\varphi \triangleright (\triangleright_{\alpha_1} \psi))$

REM(r)

14) $(\phi \triangleright_{\alpha_1} \psi) \triangleright (\varphi \triangleright (\triangleright_{\alpha_1} \psi))$

Simplification

□

Conclusion

I. Explanatory Tasks	201
II. Logical Expressivism: Challenges and Applications	203
II.I. Challenges	204
II.II. Applications	205

I. Explanatory Tasks

Now that we've formulated an expressivist view of the logical constants which covers both the connectives and the quantifiers, we've completed the explanatory tasks set out in the introduction. The first task was to explain why the logical vocabulary is there at all. The rival explanation was a metaphysical one. According to that explanation, the logical vocabulary is there for much the same reason that the physical vocabulary is there: to help us describe an aspect of the world we live in. On this view, there is logical structure in reality, and we need to get at it. The logical constants are our devices for doing so.

On the expressivist story, the logical constants have a different function. According to expressivism, we are creatures that have both simple and complex propositional attitudes. We have a need to communicate these to each other, in order to collectively get by in this world. Therefore, we need a language that we can use to express both simple and complex propositional attitudes. It's the presence of the logical constants in our language that makes our language suitable for this purpose. With these devices, we can form linguistic vehicles for the expression of propositional attitudes of any complexity.

The second task was to explain why the logical constants are semantically determinate to the degree that we think they are. The explanation I gave made use

of the ideology of expressivism. We are creatures that have a rich array of simple and complex propositional attitudes. An ideal interpreter of us would assume, from the get-go, that we have such attitudes, and that our use of language is in large part an effort to communicate these. Therefore, it is incumbent on the ideal interpreter to make our language come out as one that is useful for this purpose. It must be language that has devices for forming expressions for complex propositional attitudes. Those devices in the language the use of which best fit this structural role are attributed 'merely expressive' meanings – a mechanism which I called 'reverse reference magnetism'. These are the logical constants.

The reason that the logical constants, so interpreted, have determinate meanings is that their interpretations are derived from notions that the interpreter, qua interpreter, has at their disposal. The interpreter has notions like accepting, rejecting, ruling in and ruling out, because these notions are part of the basic psychology that the interpreter projects onto the interpretees. If you will, there are terms for these attitudes in the interpreter's own vocabulary. We assume that they are determinate in that vocabulary, for if we are not allowed to assume that the interpreter has a semantically determinate vocabulary, it would be hard to see how the very problem of radical indeterminacy could be stated. But if one has doubts on that score, it is possible to bring in a hypothesis of naturalness at this stage: that the attitudes of acceptance, rejection, etc. are (relatively) natural.

When this metasemantic proposal was presented in chapter IV, Sequentese had not yet been extended to the quantifiers. The introduction of the quantifiers required us to complicate the story in a few ways, for instance by introducing new attitudes towards objects and concepts, and the notion of a condition. But since the metasemantics presented in chapter IV were pitched at a high level of generality, these new additions to the ideology of Sequentese do not, I believe, present any additional metasemantic trouble. The ideal interpreter ought to be regarded as working on the assumptions that the language-users they interpret have the full set of attitudes that logical expressivism attributes to them, including the ones that are involved in quantification.

By addressing these two explanatory challenges, I have intended to undermine

some reasons that Ted Sider has presented for believing in logico-metaphysical structure: his arguments from semantic determinacy and ideological commitment. If I have succeeded in that, this doesn't show that there isn't logico-metaphysical structure is a theoretical posit and a new addition to our ideology, I take it that we require positive reasons to believe in it. Without such reasons, by believing it we would just clutter up our worldview.

But if I have made a convincing case against Sider's logical realism, it should be clear that this leaves other questions about the metaphysics of logic entirely unanswered. We've said nothing about the ontology of logic, for instance. In setting out logical expressivism, I've helped myself to set-theoretical notions here and there. I do not take this to commit the logical expressivist to the existence of sets definitively. There is a whole separate and more general debate to be had about the metaphysical significance of set-theoretical talk, and I take this just to be a special case of that debate. We should therefore wait until the philosophers of mathematics have made up their mind about the general question, before we decide on the ontology that logical expressivism commits us to.

Other questions have been addressed to some degree. Does logic have a normative significance? On the expressivist story, it does. In chapter III, logical implications came out as a kind of hypothetical rational permissions to have certain attitudes, accompanied by hypothetical rational impermissions to have certain other attitudes. The normative import of logical consequence is thus traced back to the more general normativity that attaches to theoretical rationality. There is a metaphysical question to be asked about that sort of normativity, of course. But like the question of the metaphysical commitments of set-theoretic talk, it's a question that goes beyond the remit of this thesis.

Despite the limited ambitions of this thesis, logical expressivism is nevertheless a theory that may have other uses in the philosophy of logic beyond the ones I've put it to. Let's discuss the features and prospects of the view a bit more.

II. Logical Expressivism: Challenges and Applications

II.I. Challenges

Logical expressivism, insofar as I have set it out, covers only the ‘core’ logical vocabulary. These are the terms that often get marked out, by those engaged in the debate on logical constant-hood, as the terms that are plausibly regarded as purely logical. I’ve not tried to contribute to that debate. It would be interesting if it turned out that one can treat these terms, and only these terms, as merely expressive in the sense I’ve explained. But if it doesn’t turn out that way, this does not seem particularly problematic.

For logical expressivism to be viable, it is not necessary that we be able to tell an expressivist story about other logical vocabulary: modal, second-order, temporal, what have you. It may turn out that these terms have a metaphysical significance that the connectives and quantifiers don’t. But it *is* important that the logical notions that get an expressivist treatment be able to interact and co-exist with these other notions. If it turned out, for instance, that the Sequentese formalism could not handle the introduction of modal or second-order vocabulary (regardless of whether this vocabulary then gets an expressivist treatment) that would be a bad result. So there is a task there: to at least show that the Sequentese formalism can be extended to some of these bits of extra logical vocabulary.

I am not too pessimistic about this task. With regard to modal operators, for instance, I think there may be prospects for not only extending the Sequentese formal logic to include them, modal operators, but also extending the expressivist semantics to cover them. Expressivism about modal notions is already an existing research programme: it will be interesting to see how the logical expressivist project might interact with that research programme and perhaps borrow from it.¹³⁶ With regard to second-order vocabulary, I suspect the extensions may be fairly straightforward.

136. See Blackburn (1993, chapter 3) and Divers & Elstein (2012). It should be noted that the latter are explicitly neutral on the realist or anti-realist implications of the view, though they note the usefulness of their view to anti-realist views of modality.

Another dimension along which logical expressivism can hopefully be developed is the treatment of non-classical logics. Classical logic was taken as the target logic in this thesis, for two reasons: it is a widely accepted logic which enjoys a certain 'default' status, and it is a relatively simple logic. I am not, however, particularly wedded to it. It would be interesting to see whether some standard intuitionist and paraconsistent logics could be reproduced in the expressivist manner. This does not appear as if it is going to be a very straightforward process. In setting up classical expressivism, I let the attitude of rejection behave much like the acceptance of a negation. But paraconsistent logicians treat rejection as stronger than the acceptance of a negation (one can accept $\neg A$ without rejecting A) and intuitionist logicians treat it as weaker (one can reject A without accepting $\neg A$). This looks to make things considerably more complicated.

I am inclined to think, however, that the prospects for intuitionist expressivism look somewhat better than those of paraconsistent expressivism. We already saw that if we gave up the principles AA and AAA, we would not get the law of excluded middle coming out as a theorem of Sequentese. Unfortunately, it looked as if, without AA and AAA, we also would not get the rule of Transitivity. But it is not inconceivable that there are some principles of rationality to be found that are weaker than AA and AAA but still strong enough to get something like Transitivity, and then we'd be on the way to developing a paracomplete logic.

II.II. Applications

Logical expressivism will hopefully have some bearing on questions in the philosophy of logic that fall outside of metaphysics and metasemantics. One interesting area of application may be the epistemology of logic, more specifically the epistemology of basic inference rules. There is, at present, a lively debate on these issues, much of it to do with the type of regress described by Lewis Carroll in his (1985).

In chapter III, the inference rules of the Sequentese formalism were derived from

principles of synchronic rationality. This suggests an interesting direction of research. How do we know such principles? Can we come to know them without the use of discursive reasoning, i.e. without the application of inference rules? How could we know the validity of inference rules on the basis of principles of rationality? Would that have to involve discursive reasoning? I am keen to explore these issues in the future.

Another, perhaps surprising area of application is 'local' expressivisms, that is, expressivisms about particular areas of discourse. There is, of course, a large research programme devoted to developing an expressivist theory of moral discourse. Logical expressivism can provide some useful resources for this research programme, as I'll try to show now.¹³⁷

Local expressivism has traditionally been plagued by a set of issues grouped together as the 'Frege-Geach problem'.¹³⁸ Local expressivisms face the challenge of providing a semantics for certain terms (moral terms, say) which is non-representational, but makes these terms behave much like they were representational. We seem to be able to do complicated reasoning using moral terms and non-moral terms together. Moral claims can embed in conditionals, under negations and in all sorts of other situations. But if moral claims, at the end of the day, aren't descriptive statements, some of this behaviour becomes hard to explain. This is sometimes brought out by means of a toy expressivist semantics, which equates the force of saying "x is good" with that of exclaiming "hooray for x!" Using that toy semantics, we can make the problem vivid. Take the following two sentences:

1. "If alms-giving is good, I should do it more."
2. "If hooray: alms-giving, I should do it more."

The first sentence makes perfect sense, the second doesn't. But "alms-giving is good" is supposed to have the same force as "hooray: alms-giving", so the one ought to substitute in for the other with preservation of sense.

137. For an overview, see Schroeder (2010).

138. Frege (1919) and Geach (1963).

This is one of the very simplest examples of a Frege-Geach problem. Moral expressivists have since tried to calibrate their semantics for moral vocabulary to stop these bad results from occurring, but the problem has a habit of re-occurring in some more complicated form even for more complicated and subtle expressivist proposals.¹³⁹

Problems of the Frege-Geach sort do not occur for logical expressivism, not because the view is so much cleverer but because the aims of the theory are different. Local expressivisms face the challenge of integrating a vocabulary which has a representationalist semantics with one that has an expressivist semantics, in such a way that the statements that can be formed with those vocabularies are logically on a par. Wherever a descriptive statement can occur, a moral claim should be able to occur, and vice versa. This turns out to be difficult.

Things are different with logical expressivism. We are indeed integrating two vocabularies with different sorts of semantics: a non-logical vocabulary with a representational semantics and a logical vocabulary with an expressivist semantics. But we do not need to make the logical terms behave like the non-logical terms: their behaviour was entirely different to begin with. The logical vocabulary is, all in all, a particularly easy target for expressivist treatment.

However, with the logical expressivist framework at our disposal, it seems we can help the local expressivist out a little, simply by enriching our stock of basic attitudes a little. I'll illustrate this by taking moral expressivism as a case study.

Suppose that we introduce a new attitude to objects, *forbidding*. Let's say that forbidding is what one does to things to which one objects morally. This is an attitude which one presumably most appropriately bears to acts, so let's restrict the possible objects of the attitude to just the acts. The act of forbidding will be formally entirely analogous, in the expressivist framework, to the act of applying a concept. That means it will be the sort of thing we can accept or reject. So we then have two possible 'uses' of forbidding: the acceptance of a forbidding and the rejection thereof. These attitudes of accepting and rejecting a forbidding can then

139. Again, see Schroeder (2010)

be the object of rulings-in and rulings-out, just like attitudes of accepting and rejecting applications of concepts.

Is it okay to just stipulate that forbidding is entirely like applying a concept when it comes to interacting with other attitudes? But to do so does not seem to do violence to the notion of forbidding, understood as moral censure. It does seem to behave like applying a predicate in these respects. For instance, the attitudes of accepting and rejecting forbiddings can presumably be exhaustively distributed over the objects. All acts are either appropriate to forbid, or they're not appropriate to forbid. So the forbidding of them can always either be accepted or rejected (once all the evidence is in). Their distribution presumably also ought to be exclusive. No act will be appropriate to forbid and not forbid. So in no case should one accept and reject the forbidding of the same act.

When it comes to building the notion of forbidding into Sequentese (which I won't do here in any detail) it will slot in where the notion of applying a concept does. So let's say, for the sake of argument, that " $\nabla|x$ " signifies the forbidding of x. Then by putting this phrase on either side of a pointer, we can signify the acceptance or rejection of that forbidding. For example:

" $\nabla|x \blacktriangleright$ "

This would signify the ruling out of the acceptance of the forbidding of x. If we put it on the other side of the pointer, it would signify the ruling out of the rejection of the forbidding of x.

Now let's do some simple semantics. First, let's make the negative claim there is no concept of moral wrongness (or rightness) that one can apply to objects, or at least not one that objects satisfy in virtue of having a corresponding property. That makes our semantics for wrongness non-representational. Second, let's make the following positive claim: that an atomic moral claim of the form "act x is wrong" expresses that one rules out the rejection of the forbidding of x. This is entirely analogous to the way that an atomic non-moral claim like "Fido is flea-bitten" expresses that one rules out the rejection of the application of the concept of flea-bitness to Fido.

We now have an important fragment of the moral semantics. We need to do two things. First, we need to either find ways of casting other moral talk in terms of forbidding, or we need to paraphrase other moral talk in terms of the predicate “is wrong”. The latter seems the easiest. For instance, we can say that something is right iff it is not wrong. I’m just going to assume, for the sake of argument, that we can paraphrase all the talk that we need to paraphrase in terms of wrongness. I’m sure it would be a tricky endeavour, but it’s beyond the scope of this thesis.

Second, we need to explain how logically complex moral talk gets analysed. Here is where the Frege-Geach problem would kick in. But here is also where we need to do the least work: for here the machinery of logical expressivism kicks in. Take the mixed moral/non-moral conditional claim “If Joe's killing was a murder, it was wrong”. The attitude which this expresses can be the following:

“The attitude of ruling out accepting the application of being a murder to Joe's killing, while rejecting the forbidding of Joe's murder”.

Or let’s look at a universal moral claim: “Alms-giving is right”. This expresses the following attitude:

“The attitude of ruling out, for all x , accepting that x is a case of alms-giving and accepting the forbidding of x .”

The process of coming up with these translations isn’t in any way novel. If our semantics for “is wrong” were in terms of the application of a concept of wrongness, we’d have come up with similarly structured translations. It’s just the translation of the “is wrong” bit that will differ. Adjusting the translation manuals in Chapters III and V will, not, therefore, be complicated.

I won't go and prove it in detail here, but it should be apparent that by building in the moral 'predications' at the bottom level of the theory, we avoid the task of guaranteeing separately that the new attitudes that we've introduced for the purpose behave properly at every level. We are guaranteed, in effect, that a claim of moral wrongness, despite not being analysed as a predication, can occur in any place where an atomic predication could, with the same logical role. In this way we can nip Frege-Geach issues in the bud.

This semantic approach is compatible with both realist and anti-realist forms of expressivism. If one wants to maintain that moral claims can be correct or incorrect, then one can do this by building appropriate clauses into the inductive definition of correctness given in chapter V.

This is only the barest sketch of an expressivist theory of morality. It is not my business here to defend any kind of local expressivism. But I hope to have shown that there are some promising applications for logical expressivism in this area.

Bibliography

- Akiba, K., 1996. Logic as instrument: the Millian View on the Role of Logic. *History and Philosophy of Logic* 17:1-2, pp. 73-83.
- Beall, JC & Restall, G., 2005. *Logical Pluralism*. Oxford: Clarendon Press.
- Belnap, N.D., 1962. Tonk, Plonk and Plink. *Analysis* 22(6), pp. 130-134.
- Bendall, K., 1979. Negation as a Sign of Negative Judgment. *Notre Dame Journal of Philosophical Logic* 20(1), pp. 68-76.
- Blackburn, S., 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- Boghossian, P.A., 1989. The Rule-Following Considerations. *Mind* 98(392), pp. 507-549.
- Boghossian, P., 2003. Blind Reasoning. *Proceedings of the Aristotelian Society*, Supplementary Volume 77, pp. 225–248.
- Burge, T., 1979. Individualism and the Mental. *Midwest Studies in Philosophy* 4(1), pp. 73-121.
- Carnap, R., 1950. Empiricism, Semantics, and Ontology. *Revue Internationale de Philosophie* 4(2), pp. 20-40.
- Carroll, L., 1895. What The Tortoise Said To Achilles. *Mind* 4(14), pp. 278-280.
- Chalmers, D.J., 2012. *Constructing the World*. Oxford: Oxford University Press.
- Divers, J. & Elstein, D., 2012. Manifesting Belief in Absolute Necessity. *Philosophical Studies* 158(1), pp. 109-130.
- Field, H.H., 1975. Conventionalism and Instrumentalism in Semantics. *Nous* 9 (4), pp. 375-405.

- Field, H.H., 1980. *Science without Numbers*. Princeton: Princeton University Press.
- Field, H.H., 2009. What is the Normative Role of Logic. *Proceedings of the Aristotelian Society*, Suppl. Vol. 83, pp. 251–68.
- Frege, G., 1919. Die Verneinung. Republished in: Patzig, G., ed. 2003 (5th edition). *Logische Untersuchungen*. Göttingen: Vandenhoeck & Ruprecht, pp. 63-83.
- Geach, P. T., 1963. Assertion. Reprinted in: Geach, P. T., 1972. *Logic Matters*. Oxford: Blackwell, pp. 254-269.
- Gibbard, A., 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Clarendon Press.
- Goodman, N., 1979. *Fact, Fiction and Forecast*. Cambridge, MA: Harvard University Press.
- Hattiangadi, A., 2006. Is Meaning Normative? *Mind and Language* 21(2), pp. 220-240.
- Hawthorne, J., 2007. Crazy and Metasemantics. *Philosophical Review* 116(3), pp. 427-440.
- Hirsch, E., 2002. Quantifier Variance and Realism. *Philosophical Issues* 12 (Realism and Relativism), pp. 51-73.
- Hirsch, E., 2004. Comments on Theodore Sider's *Four Dimensionalism*. *Philosophy and Phenomenological Research* 68(3), pp. 658-664.
- Hirsch, E., 2005. Physical-Object Ontology, Verbal Disputes, and Common Sense. *Philosophy and Phenomenological Research* 70(1), pp. 67-97.
- Hirsch, E., 2008. Language, Ontology and Structure. *Nous* 43 (3), pp. 509-528.
- Hirsch, E., 2009. Ontology and Alternative Languages. In: Chalmers, D., Manley, D., Wasserman, R., eds. 2009. *Meta-metaphysics*. Oxford: Oxford University Press, pp. 231-259.

- Kripke, S.A., 1972. *Naming and Necessity*. Oxford: Blackwell.
- Kripke, S., 1982. *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- Lewis, D.K., 1969. *Convention*. Cambridge: Cambridge University Press.
- Lewis, D.K., 1974. Radical Interpretation. *Synthese* 27(3), pp. 331-344.
- Lewis, D.K., 1975. Languages and language. In: Gunderson, K., ed. 1975. *Minnesota Studies in the Philosophy of Science* 7. Minneapolis: University of Minnesota Press, pp. 3-35.
- Lewis, D.K., 1984. Putnam's Paradox. *Australasian Journal of Philosophy* 62(3), pp. 221-236.
- Lewis, D.K., 1992. Meaning without Use: Reply to Hawthorne. *Australasian Journal of Philosophy* 70(1), pp. 106-110.
- Lewis, D.K., 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61(4), pp. 343-377.
- MacFarlane, J., (unpublished). In What Sense (If Any) Is Logic Normative for Thought? Available at: <<http://johnmacfarlane.net>> [Accessed 18/01/12].
- Manley, D., 2009. When Best Theories Go Bad. *Philosophy and Phenomenological Research* 78(2), pp. 392-405.
- Marsh, G., 2009. Is the Hirsch-Sider Dispute Merely Verbal? *Australasian Journal of Philosophy* 88(3), pp. 459-469.
- McDaniel, K., 2009. Ways of Being. In: Chalmers, D., Manley, D., Wasserman, R., eds. 2009. *Meta-metaphysics*. Oxford: Oxford University Press, pp. 290-319.
- McGee, V., 2006. There's a Rule for Everything. In: Rayo, A & Uzquiano, G., eds. 2006. *Absolute Generality*. Oxford: Oxford University Press, pp. 179-202.

- McGrath, M., 2008. Conciliatory Metaontology and the Vindication of Common Sense. *Nous* 42(3), pp. 494-495.
- Melia, J., 1995. On What There's Not. *Analysis* 55(4), pp. 223-229.
- Peacocke, C.A.B., 1992. *A Study of Concepts*. Cambridge, MA: MIT Press.
- Price, H., 1990. Why 'Not'? *Mind* 99(394), pp. 221-238.
- Price, H., (unpublished). 'Not' Again. Available at <<http://prce.hu>> [Accessed 18/01/12].
- Priest, G., 2005. *Doubt Truth to be a Liar*. Oxford: Oxford University Press.
- Prior, A.N., 1960. The Runabout Inference Ticket. *Analysis* 21(2), pp. 38-39.
- Putnam, H., 1975. The Meaning of "Meaning". In: Gunderson, K., ed. 1975. *Minnesota Studies in the Philosophy of Science* 7. Minneapolis: University of Minnesota Press, pp. 131-193.
- Putnam, H., 1980. Models and Reality. *Journal of Symbolic Logic* 45(3), pp. 464-482.
- Quine, W.V.O., 1936. Truth by Convention. In: Lee, O.H., ed. 1936. *Philosophical Essays for A.N. Whitehead*. New York: Longmans, pp. 90-124.
- Quine, W.V.O., 1948. On What There Is. *Review of Metaphysics* 2(5), pp. 21-36.
- Quine, W.V.O., 1951. Ontology and Ideology. *Philosophical Studies* 2(1), pp. 11-15.
- Ramsey, P.F., 1929. General Propositions and Causality. In: Mellor, D.H., ed. 1990. *P.F. Ramsey: Philosophical Papers*. Cambridge: Cambridge University Press, pp. 133-151.
- Resnik, M.D., 1999. Against Logical Realism. *History and Philosophy of Logic* 20(3-4), pp. 181-194.
- Restall, G., 2005. Multiple Conclusions. In: Hajek, P., Valdes-Villanueva, L.,

Westerstahl, D. eds. 2005. *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*. London: Kings' College Publications, pp. 189-205.

Restall, G., 2008. Assertion and Denial, Commitment and Entitlement, and Incompatibility (and some consequence). *Studies in Logic* 1, 26–36.

Restall, G. (unpublished). Assertion, Denial and Non-classical Theories. Available at: <<http://consequently.org/writing>> [Accessed on 18/01/12].

Ripley, D., (forthcoming(a)). Conservatively Extending Classical Logic With Transparent Truth. Forthcoming in *Review of Symbolic Logic*.

Ripley, D., (forthcoming(b)). Paradoxes and Failures of Cut. Forthcoming in *Australasian Journal of Philosophy*.

Rumfitt, I., 2000. 'Yes' and 'No'. *Mind* 109(463), pp. 781-824.

Schroeder, M., 2010. *Noncognitivism in Ethics*. London: Routledge.

Schwartz, W., (2012 draft). Lewisian Meaning without Naturalness. Available from <www.umsu.de/papers> [Accessed 17-01-13].

Sider, T., 2001a. *Four-Dimensionalism*. Oxford: Oxford University Press.

Sider, T., 2001b. Criteria of Personal Identity and the Limits of Conceptual Analysis. *Philosophical Perspectives* 15, pp. 189-209.

Sider, T., 2004. Replies to Gallois, Hirsch and Markosian. *Philosophy and Phenomenological Research* 68(3), pp. 674-687.

Sider, T., 2009a. Ontological Realism. In: Chalmers, D., Manley, D., Wasserman, R., eds. 2009. *Meta-metaphysics*. Oxford: Oxford University Press, pp. 384-423.

Sider, T. 2009b. Against Vague and Unnatural Existence: Reply to Liebesman and Eklund. *Nous* 43(3), pp. 647-687.

Sider, T. 2013. "Against Parthood". Forthcoming in: Bennett, K. & Zimmerman, D.W., *Oxford Studies in Metaphysics*, Vol. 8. Oxford: Oxford University Press. Available at <<http://tedsider.org>> [Accessed 23/04/12].

Sider, T., 2010. *Logic for Philosophy*. Oxford: Oxford University Press.

Sider, T., 2011. *Writing the Book of the World*. Oxford: Oxford University Press.

Smiley, T., 1996. Rejection. *Analysis* 56(1), pp. 1-9.

Turner, J.T. 2010. Ontological Pluralism. *Journal of Philosophy* 107(1), pp. 5-35.

Weatherson, B., 2003. What Good Are Counterexamples? *Philosophical Studies* 115(1), pp. 1-31.

Weatherson, B., (unpublished). The Role of Naturalness in Lewis's Theory of Meaning. Available at <brian.weatherson.org> [Accessed 17-01-13].

Westerståhl, D., 2001. Quantifiers. In: Goble, L, ed. 2001. *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell, pp. 437-460.

Williams, J.R.G., 2007. Inscrutability and Eligibility. *Philosophical Review* 116(3), pp. 361-399.

Williams, J.R.G., (unpublished). Reference Magnets. Available at <robert.williams.org> [Accessed 17/01/13].

Wittgenstein, L., 1921. *Tractatus Logico-Philosophicus* (Logisch-Philosophische Abhandlung). In: Wilhelm Ostwald, W., ed. *Annalen der Naturphilosophie* 14.

Wittgenstein, L., 1953. *Philosophische Untersuchungen*. Oxford: Blackwell.