Absolute and Relative Generality

A thesis submitted for the degree of Doctor of Philosophy in Philosophy

James Studd

Michaelmas Term 2011
Corpus Christi College
University of Oxford

Acknowledgements

This thesis has greatly benefitted from the generous help of others. Thanks to audiences at ENS in Paris, Birkbeck in London, and Oxford where parts of it were presented, especially to Sam Roberts and Denis Bonnay for helpful written comments. Thanks to my colleagues, Volker Halbach, Philip Welch, Kentaro Fujimoto, Graham Leigh and Carlo Nicolai, for a great deal of valuable input and welcome encouragement. Thanks to my examiners, Tim Williamson and Øystein Linnebo, for their careful reading, incisive comments and searching examination. Thanks especially to my second supervisor, Alex Paseau, for guiding me safely through the final stages, leading to a great many improvements in clarity along the way. Thanks most especially to my first supervisor, Gabriel Uzquiano, for his inspirational tuition, which played a vital role in the genesis and flourishing of many of the ideas in this thesis. I should also like to gratefully acknowledge the financial support of the AHRC during the first seven terms in which this thesis was written.

Abstract

Absolute and Relative Generality

A thesis submitted for the degree of Doctor of Philosophy in Philosophy

James Studd, Corpus Christi College, the University of Oxford

Michaelmas term 2011

Word count: 74549

This thesis is concerned with the debate between absolutists and relativists about generality. Absolutists about quantification contend that we can quantify over *absolutely* everything; relativists deny this. The introduction motivates and elucidates the dispute. More familiar, restrictionist versions of relativism, according to which the range of quantifiers is always subject to restriction, are distinguished from the view defended in this thesis, an expansionist version of relativism, according to which the range of quantifiers is always open to expansion.

The remainder of the thesis is split into three parts. Part I focuses on generality. Chapter 2 is concerned with the semantics of quantifiers. Unlike the restrictionist, the expansionist need not disagree with the absolutist about the semantics of quantifier domain restriction. It is argued that the threat of a certain form of semantic pessimism, used as an objection against restrictionism, also arises, in some cases, for absolutism, but is avoided by expansionism. Chapter 3 is primarily engaged in a defensive project, responding to a number of objections in the literature: the objection that the relativist is unable to coherently state her view, the objection that absolute generality is needed in logic and philosophy, and the objection that relativism is unable to accommodate 'kind generalisations'. To meet these objections, suitable schematic and modal resources are introduced and relativism is given a precise formulation.

Part II concerns issues in the philosophy of mathematics pertinent to the absolutism/relativism debate. Chapter 4 draws on the modal and schematic resources introduced in the previous chapter to regiment and generalise the key argument for relativism based on the set-theoretic paradoxes. Chapter 5 argues that relativism permits a natural motivation for Zermelo-Fraenkel set theory. A new, bi-modal axiomatisation of the iterative conception of set is presented. It is argued that such a theory improves on both its non-modal and modal rivals.

Part III aims to meet a thus far unfulfilled explanatory burden facing expansionist relativism. The final chapter draws on principles from metasemantics to offer a positive account of how universes of discourse may be expanded, and assesses the prospects for a novel argument for relativism on this basis.

Contents

1 Introduction

1 Absolutism

The absolutist about quantification contends that sometimes quantifiers such as 'everything' can be used to express universal quantification over *absolutely* everything.

1.1 Putatively absolute generality

This view enjoys a good deal of initial plausibility. Often in philosophy, science and mathematics it suffices to quantify over *less* than everything. The theory at hand may call for us to generalise only about agents with free will or perfectly natural properties; about particles in the standard model or varieties of archaea; or about Banach spaces or vorticity-free fluids; and so on. But sometimes restricted generality does not seem to be enough: to theorise with the requisite level of generality requires us to quantify over *everything.*

Take, for instance, the mereological nihilist. Having explained that simple things are those that have no proper parts, he makes the following proclamation.

(1) Everything is simple.

Well aware of the potential for such a radical claim to invite misunderstanding, the mereologist takes pains to emphasise that he does not intend his use of 'everything' to be restricted, and will happily elaborate on his use of 'everything', with utterances like: 'By 'everything', I mean to be included absolutely everything whatsoever, without any exclusions at all.'[1]

[1] Compare Williamson (2006, p. 380).

This provides a *prima facie* reason to side with the absolutist about quantification, who contends that quantification over *absolutely* everything is possible. But there is also a well known *prima facie* case against this view.

1.2 Indefinite extensibility

Our pre-theoretic understanding of certain predicates seems to rule out our pinning down a final—*in*extensible—extension for them. Predicates such as 'set', 'ordinal', 'interpretation', amongst others, appear to be 'indefinitely extensible' in something like Michael Dummett's sense:[2] no matter what sets a putative extension for 'set' comprises, we can apparently specify a further set that is not one of them; similarly for 'ordinal' and 'interpretation'.

Take sets first. Intuitively a set is just a collection of some things. The things are its elements. The set of some things is identical with the set of some other things just in case the first things are the same as the second things; that is, just in case the sets have the same elements. But given any sets, reasoning in the style of Russell's (1902) paradox appears to lead us to a further set.

Consider those sets amongst the given ones that lack themselves as elements. Now consider the Russell set, which has these non-self-membered sets as its elements. Suppose for *reductio ad absurdum* that the Russell set *is* amongst the sets we started with. If the Russell set is non-self-membered, it is one of the given sets that is non-self membered, and hence one of the elements of the Russell set. If it is self-membered, since the Russell set has as elements only sets that are *non*-self-membered, it is not an element of the Russell set after all. This yields a contradiction since it *is* the Russell set. The Russell set is self-membered just in case it is not. Consequently 'set' would seem to be indefinitely extensible. Given any sets that putatively exhaust the extension of 'set', we appear to be able to come to a further set—namely, the Russell set for those sets—which cannot be amongst them.[3]

[2] See Dummett (1963; 1983, pp. 532–3; 1991, pp. 316–19). I shall offer a regimentation of this notion in chapter 5.

[3] This reasoning is the plural analogue of the set-theoretic argument employed by Zermelo to show that every set M has a subset $M_0 =_{\mathrm{df}} \{x \in M \mid x \notin x\}$ that is not an element of M. See Zermelo (1908, thm. 10).

Next take interpretations. Focus just on the interpretation of a single unary predicate P from a formal language. An interpretation specifies which things P applies to; that is, which things satisfy the open sentence Px. Intuitively P can be interpreted to apply to any things whatsoever. For example, we can interpret P to have the same extension as 'bottle' by taking it to apply to every bottle (and nothing else). Under this interpretation, the predicate does not apply to non-bottles; more specifically, P does not apply to interpretations; and, in particular, does not apply to the bottle interpretation itself. But in analogy with the Russell argument, by considering such 'non-self-applying' interpretations, we can seemingly show how to extend any putatively complete extension of 'interpretation'. The argument adapts an interpretational analogue of the Russell paradox due to Williamson (2003, §IV). Given any interpretations, consider the Williamson interpretation, which takes P to apply to the non-self-applying interpretations amongst them; that is, to the given interpretations I under which P does not apply to I. Then—supposing the Williamson interpretation is amongst the given interpretations—if it is non-self-applying, the Williamson interpretation takes P to apply to it. If it is self-applying, since the Williamson interpretation applies P only to interpretations that are *non*-self-applying, P does not apply to it under the Williamson interpretation. Exactly analogous reasoning to before leads to an exactly analogous contradiction. The Williamson interpretation is self-applying just in case it is not.

Finally take ordinals. Intuitively ordinals provide a measure of the 'length' of certain orderings: an ordinal α is just the order-type of some things, well-ordered by an associated relation $<_\alpha$.[4] The ordinal α of some well-ordered things is identical with the ordinal β of some other well-order things if the first things (under $<_\alpha$) exhibit the same order-structure as the second things (under $<_\beta$); more precisely, just in case the first things are order-isomorphic to the second things under their respective orderings.[5]

[4] A relation $<_\alpha$ is said to *well-order* some things if it is a strict total order on those things—an irreflexive, transitive and connected relation on those things—that is well-founded: such that for one or more things amongst the ordered things, one of them a is minimal, that is, such that no b amongst them is such that $b <_\alpha a$.

[5] Some things, ordered by $<_\alpha$, are said to be *order-isomorphic* to some others, well-ordered by $<_\beta$, if there is bijection π between the first things and the second things such that any two of the first things a and b stand in the first relation—$a <_\alpha b$—just in case their counterparts amongst the second things stand in their order relation—$\pi(a) <_\beta \pi(b)$.

But, as with sets, the intuitive conception of ordinal appears to preclude our settling on an intuitively exhaustive extension of 'ordinal'. For given any ordinals, the reasoning of the Burali-Forti (1897) paradox seems to take us to a further ordinal. The argument this time is slightly more involved. First we define the standard ordering on the given ordinals: one ordinal is 'shorter' than another if the first well-ordering has same order-structure as a proper part—in particular, an initial segment—of the second: $\alpha < \beta$ if α is identical to $\beta^{<b}$, the order-type of those things *prior to* b under $<_\beta$ (well-ordered under the restriction of $<_\beta$ to these things), for some b amongst them. We can then establish that $<$ well-orders the ordinals in the purportedly exhaustive extension of 'ordinal'. But now consider the Burali-Forti order-type Ω of all the given ordinals as ordered by $<$. The assumption that Ω is one of the given ordinals leads to a contradiction. For if Ω is one of these ordinals, the ordinals less than Ω are an initial segment of all the given ordinals, and $\Omega^{<\Omega} < \Omega$. Given the well-foundedness of $<$, there must be a minimal ordinal β with this property, such that $\Omega^{<\beta} < \beta$. But then the shorter ordinal $\alpha =_{\text{df}} \Omega^{<\beta}$ also has this property. The ordinals prior to α are an initial segment of the ordinals prior to β; so $\Omega^{<\alpha} < \Omega^{<\beta}\, (= \alpha)$. This contradicts the minimality of β.[6]

2 Relativism

Arguments in this vein lie behind what I take to be by far and away the strongest case for relativism about quantification, the view opposing absolutism. Relativists moved by such considerations maintain that whatever objects the absolutist's supposedly absolutely all-inclusive domain comprises, we can come to use quantifiers more extensively, so that—relative to the new, more liberal perspective—there are things that are not members of this domain.[7]

The picture is akin to the one Zermelo (1930) outlined for the set-theoretic hierarchy. Rather than a single intended model for set theory with urelements ZFCU, Zermelo hypothesised an open-ended hierarchy of set-theoretic universes, each more inclusive than the last. Some objects that do not form a set in one model, the members of the

[6]Shapiro and Wright (2006, pp. 256–7) offer a slightly different formulation of this argument.
[7]See, for instance, Glanzberg (2006, p. 50).

proper class Ω of all (von Neumann) ordinals in the universe of the model, for instance, or simply every set in the entire universe of the model, $V[U]$, form a perfectly good set in the next, more inclusive model. And the series continues indefinitely, surpassing the order-type measured by any ordinal in any of its members. Likewise the relativist contends that rather than a single absolute domain for quantifiers like 'everything', there is a series of ever greater domains. We can always bring about a shift in domain so as to incorporate the Russell set of all non-self membered sets initially quantified over, or the Williamson interpretation of the non-self-applying interpretations of P in the initial domain, or the Burali-Forti order-type of the ordinals in it. And, by the above arguments, the new domain cannot be a subdomain of the initial one.[8]

We shall not be ready to formulate this argument precisely in full generality until a much fuller account of shifts in domain is in place and certain dialectical issues have been addressed. But it is worth making some preliminary comments all the same. The relativist should not overstate the argument. The arguments from paradox do not show that absolutism is paradoxical in itself. For the arguments rely on non-logical premisses about sets, ordinals or interpretations. But nor should the absolutist understate this argument. The premisses about sets and the like need not themselves be inconsistent; in particular the 'intuitive' conception of set need not rely on the inconsistent Naïve Comprehension schema, or its plural analogue, which maintains that for any zero or more things *there is* a set with those things as its elements. For unlike the principle behind the relativist's argument, this fails to account for a potential shift of perspective. Rather, as we shall see in chapter 4 when we offer precise formulations of this argument, inconsistency arises from individually consistent assumptions, with absolutism one amongst many. But consistency does not amount to truth; and I will argue that the costs of rejecting the other assumptions are high.

In the end, neither view will be seen to be inconsistent or obviously false. Whether we should ultimately prefer absolutism or relativism will have to come down to an assessment of their relative merits. In carrying out this assessment, however, it is necessary to

[8]One domain D is a *subdomain* of another D' if every member of D is a member of D'. A subdomain D of D' is a *proper subdomain* if D' is not also a subdomain of D.

make several distinctions between types of relativism. The view admits of a wide array of variants, some considerably more plausible than others. These have often not been cleanly separated in the literature, and sometimes fare very differently when it comes to the objections that have been pressed against the generic, undifferentiated view. For now we shall make two principal distinctions, with a third following in chapter 3.

2.1 Context versus interpretation

The first decision point for the relativist is to determine what is responsible for the shifts in domain she posits. It is relatively uncontroversial that there are at least two factors that determine the domain of our quantifiers: first, how our expressions are interpreted, determining what the standing linguistic meanings of expressions like 'everything' is; and second, the context of utterance, determining their semantic content in particular utterances. This leads to at least a bifurcation in relativist views according to whether they take shifts in perspective to be due to shifts in context or shifts in interpretation.

Consider context first. Suppose that, looking at the empty bottles and glasses on the table, someone utters:

(2) Everything is empty.

In reply, taking a new bottle out of the fridge, a second speaker utters:

(3) Not everything is empty.

It seems clear that there is no disagreement here. Quantifiers are subject to contextual restrictions. What has been said, what is salient in our surroundings, and so on, often restricts the range of our quantifiers to things that are contextually relevant. The context of the first speaker's utterance restricts the quantifier 'everything' to the bottles, glasses, etc. on the table. By uttering (2), she says (roughly) that every bottle, glass, etc. *on the table* is empty. In the context of the second speaker's utterance, a more liberal contextual restriction applies—just to bottles, glasses, etc. in the room, perhaps. By uttering (3), the second speaker says that not every bottle, glass, etc. *in the room* is empty. Both sentences are true in their respective contexts: (2) is true (and (3) false) in the first, more restrictive context; (3) is true (and (2) false) in the second, wider context.

So much is fairly uncontroversial.[9] Context contributes to determining what objects a given quantifier ranges over. But such contextual relativity in quantifiers on its own is no threat to absolutism about quantification. The absolutist can happily accept that the range of our quantifiers is constrained by the context of utterance provided that some contexts leave quantifiers unrestricted. The proponent of a contextual version of relativism must maintain further that there is no such context.

The second candidate for a shift in perspective is a shift in the interpretation of the lexicon.[10] An interpretation supplies semantic values for the expressions in the language. The models that are standardly used to interpret formal first-order languages provide a first example of an interpretation. Such models supply a universe M, together with (i) a member of M as the extension of each singular term; and (ii) an n-ary relation over M as the extension of each n-ary predicate.[11] Such models allow for the quantifier $\forall$ to be reinterpreted in the manner that interests the interpretational relativist. The semantic clauses accompanying the model ensure that $\forall$ is always interpreted to express universal quantification over M. The quantifier is not open to arbitrary reinterpretation as predicates are in such models, but its domain varies from model to model. While these models are adequate for many purposes, more refined notions of an interpretation may be obtained by making allowances for richer lexicons, intensional differences in meaning, context sensitivity and so on.[12]

Shifts in interpretation have sometimes been used to explain away apparent disagreements between ontologists. To rehearse a well-worn example that originated in Putnam (1987a), imagine two linguistic communities (each of which retain the lexicon and syntax of English). The first community have long been staunch mereological ni-

---

[9] Not, however, completely uncontroversial. Bach (2000) argues that context does not affect what is said in this way but only makes a pragmatic difference. The contextual relativist must maintain, on the contrary, that quantifier domain restriction is a matter of semantics. See Stanley and Szabó (2000) for discussion and the case in favour of taking a semantic approach.

[10] Throughout we treat sentences as composed from lexical items and as having syntactic structure (as opposed to mere strings of shapes or noises). 'Sentence' always means 'declarative sentence'. In formal languages, sentences also include open sentences or formulas that have some variables free.

[11] And for languages with function symbols, an n-ary function on M as the extension of each n-ary function symbol.

[12] We return to such refinements in chapter 2.

hilists, rejecting the existence of any mereologically-complex objects. The second are devout mereological universalists, who upholds the Principle of Unrestricted Composition, maintaining that for any one or more things, something is their mereological fusion.[13] Upon meeting for the first time, member of the nihilist community attempt to convey their view to the universalists (in any context—and, in particular, even if contrary to the first sort of relativism, no contextual restrictions are in place) by uttering:

(4) Everything is simple.

Universalists (taking care to ensure that their context of utterance is not too restrictive, and if possible wholly free from restriction) rejoin with:

(5) Not everything is simple.

Some philosophers claim that there is no genuine disagreement here either.[14] Even if both utterances are made in contexts when no contextual restrictions are imposed on the range of the quantifiers, 'everything' can be interpreted in different ways. Under the nihilist's interpretation of 'everything' (and the other expressions of his language), the universe consists only of simple objects. Consequently, even if no contextual restriction is operative, in uttering (4), the nihilist says (roughly) that everything *in the nihilist's universe* is simple. By contrast, under the universalist's more liberal interpretation of 'everything', the unrestricted quantifier ranges over a universe that includes complex objects in addition to simple ones. By uttering (5), the universalist says that not everything *in the universalist's universe* is simple. Both sentences are true under their intended interpretations: (4) is true (and (5) false) under the nihilist's interpretation; (5) is true (and (4) false) under the universalist's interpretation.

Here too interpretational relativity on its own need not be highly controversial. Evidently, on the assumption that there are non-simple objects, some English quantifiers like 'every simple' are interpreted not to range over everything (in any context), all we need for interpretational relativity of the relevant kind is the possibility that some linguistic communities should only possess quantifiers with such limited interpretations (with other expressions likewise restricted).

[13]See, e.g., Lewis (1991, 73-4).

[14]See, for instance, Putnam(1987a, 1987b) and Hirsch (2002, 2005, 2009).

2.2 The argument from metaontology

Considerations along these lines have sometimes been used as a direct argument for relativism, rather than merely a means of motivating a variant of it. It is a disputed question in metaontology whether something like the view Putnam (1987a) labels 'Metaphysical Realism' holds: whether there is some language—or ontological framework, or conceptual scheme, or similar—that is 'metaphysically distinguished'. Granted a suitable elucidation of the *prima facie* obscure notion of being 'metaphysically distinguished', the thought is that, even if each side's theory is true in their own language, or framework, etc., just one mereological theory—whether nihilism or universalism (or some intermediate position)—attains the supposedly more desirable standard of being true from the metaphysically distinguished perspective. The worry, pressed by Charles Parsons (2006, pp. 205–7) and Geoffrey Hellman (2006, pp. 83–8) is that the absolutist is committed to siding with the Metaphysical Realist. As Hellman puts it: 'The absolutist must insist that... at most one of the frameworks is correct, the one (if any) that quantifies over only those objects in the range of the absolute quantifiers, the objects that 'really exist'... ' (p. 87).

There are two reasons why I do not think that this is a particularly effective attack on absolutism. The first is that I think many absolutists will find Metaphysical Realism quite congenial. They will be happy to insist, for instance, that whether or not there are mereologically complex objects is a matter fully determined by how the world is; and if, for instance, it turns out that there are complex objects, so much the worse for languages like the nihilist's that ignore them.[15]

Second, and more importantly, there is no direct link between absolutism and so-called Metaphysical Realism. As with contextual relativity, interpretational relativity *on its own* is no threat to absolutism. The absolutist can happily accept that the range of our quantifiers is constrained by the interpretation of our language, provided that under some interpretation—in particular, the absolutist will typically add, under the *actual* interpretation—'everything' encompasses absolutely everything. The absolutist,

[15] Compare Richard Cartwright in response to Dummett on whether there are sets answering to our intuitive conception of set. (Cartwright 1994, p. 15).

then, is committed to a *maximally extensive* interpretation of the quantifiers, but this is not obviously the same thing as his being committed to a 'metaphysically distinguished' interpretation. And in fact under the elucidations of this notion presented in the literature the two notions come apart. It is an irony that in their dispute over whether or not ontologists equivocate on 'something', metaontologists have likewise tended to equivocate on what they mean by 'metaphysically distinguished'. Putnam (1987a) in his attack on Metaphysical Realism, and Eli Hirsch (2002) in his defence of 'quantifier variance' have taken metaphysical distinction to amount to *expressive power*. To be 'metaphysically distinguished' a language must be able to express all the facts; where Hirsch contends moreover that facts here are to be individuated in a coarse-grained way, so that necessarily equivalent sentences express the same fact. (See Putnam (1987a, pp. 32–3); Hirsch (2002, pp. 59–62).) Meanwhile, the opposing side, prominently represented by Theodore Sider, takes 'metaphysical distinction' to be a primitive notion. The view builds on David Lewis's contention that some properties are more natural than others. *Inter alia* naturalness is hypothesised to underlie objective similarity. Since *green* is a more natural property than *grue*, two green objects are objectively more similar than two grue objects of different colours, *ceteris paribus*. (See Lewis (1983).) Sider extends such brute naturalness to quantifier denotations as well as predicate denotations, contending that some interpretations of quantifiers 'carve nature at the joints' better than others. (See Sider (2009).) But neither take on metaphysical distinction privileges more extensive interpretations of quantifiers. Sider accepts that many interpretations are available (2009, p. 392), but maintains that the nihilist's less extensive interpretation is more fundamental due to its greater ideological simplicity (2010). Hirsch maintains that the nihilist's language has just the same coarse-grained expressive power as the universalist's more liberal language (2002, p. 57).

Still, even if metaontological considerations alone do not provide support for interpretational versions of relativism, other considerations, in particular those based on indefinite extensibility, may still do so.

2.3 A third parameter?

We have at least two variants of relativism according to whether shifts in the range of our quantifiers are due to shifts in context or shifts in interpretation. Are there are any other parameters which might induce such shifts?

One factor that does affect the range of our quantifiers but is *not* plausibly taken to be behind shifts of range of the sort that interest the relativist are the circumstances: how things are with the world. Consider, for instance, a varying-domain version of presentism, which maintains that everything is present and rejects the Barcan formulas for tense operators: contending that there have been and will be things that do not exist now; and that there are things now that will not exist in the future or did not exist in the past. This view contends that the range of our quantifiers shifts as the circumstances do, from instant to instant, but it seems clear that this view does not amount to a version of relativism about quantification in an interesting sense. The absolutist claims only to quantify over absolutely everything there in fact is. It is a further claim, contra varying-domain presentism,[16] that this comprises everything that will or did exist. Likewise, one can be an absolutist while endorsing actualism and rejecting the Barcan formulas for metaphysical modal operators.[17]

Nevertheless, recent work of Kit Fine's (2005, 2006, 2007) on what he calls 'procedural postulationism' attempts to find room for a third non-circumstantial parameter. He claims that the shift in perspective involved in the Russell argument is achieved by carrying out a 'procedural postulate': 'Associated with the condition $\forall y(y \in x)$ will be an instruction or 'procedural postulate', $!x(\forall y \in x)$, requiring us to introduce an object x whose members are the objects y of the given domain'—given that no set is self-membered this amounts to an instruction to introduce the Russell set for the given domain—'. . . the intent is that there is no more fundamental understanding of what the new domain should be except as the domain that might be reached from the given domain by adding an object in conformity with the condition' (2006, p. 37). Fine is quite

[16]Although not contrary to *constant domain* presentism, which contends that despite everything being present, nothing can be brought into being or extinguished from it.
[17]This point made in this paragraph is made by Williamson (2003, p. 432).

clear that this is not an instruction to literally bring something new into existence, but he also explicitly disavows its consisting in a shift in semantic content.[18]

> There are many different ontologies, all of which have the same right (or perhaps we should say no right) to be regarded as the sum total of what there is. But this means that there is now a new way in which a statement may change its truth-value—*not through a change of content or circumstance*, but through a change in the ontology under consideration. There is *another parameter* in the picture and hence another possibility for determining how a statement may be true. Postulation then serves to fix the value of this parameter; rather than altering how things are within a given ontology or imposing a different demand on the ontology, it induces a shift in the ontology itself. (Fine 2006, p. 40, my emphasis.)

Having raised the possibility of a third parameter, I shall have little to say about it in the following. It seems to me that we have a robust pre-theoretic grasp on what it is to change the circumstances, or to shift the context of utterance, or to reinterpret our lexicon. Even if we do not always know what the mechanisms behind such shifts are, it is clear that they happen. But this is not the case for Fine's supposed third-parameter. His elucidation of postulation is set out in terms of carrying out instructions for introducing objects into the domain. But without an independent grasp of how to shift the domain without shifting the circumstances or context or interpretation, what this amounts to remains mysterious. Another aspect of Fine's view, however, will occupy us a great deal.

2.4 Restriction versus expansion

Fine distinguishes two ways in which we might reinterpret a quantifier. The first is through *de-restriction.* On this model 'the interpretation of the quantifier is given by something like a predicate or property which serves to restrict its range'. We increase the domain by relaxing the restricting condition. (2006, p. 35). For example, we might de-restrict the domain of 'every bottle' by reinterpreting the restricting predicate 'bottle' to apply not just to bottles but also to other things made of glass, say. Notice however that de-restricting the quantifier in this way to shift from the domain comprising every

[18] In his 2005 paper, Fine invokes the narrative device of a genie who carries out the postulates for us. The genie is conservative and economical. He does not make changes to previously postulated objects, and introduces as few objects as possible to conform to the command.

bottle to the new, wider domain of glass objects, relies on our having some sort of grasp on a third domain encompassing both domains. It relies on an understanding of the determiner 'every', which when combined with a universally-applicable predicate like 'thing' and subject to no other restrictions ranges over the encompassing domain. The second sort of reinterpretation does not rely on such an understanding. Rather, on the basis of understanding quantification over an initial domain, reinterpretation by *expansion* permits us to come to understand quantification over a wider domain, without presupposing some sort of grasp of a third domain encompassing both. As Fine puts it, 'We might say that the new domain is understood from 'above' under ... the restrictionist account, insofar as it is understood as the restriction of a possibly broader domain, but that it is understood from 'below' under the expansionist account, in that it is understood as the expansion of a possibly narrower domain' (p. 38).

Fine makes this distinction in the context of defending his third-parameter version of relativism, procedural postulationism, but I would like to separate the two. The distinction between restriction and expansion has wider significance in the debate about absolute generality. To further elucidate the distinction let us distinguish *domains* from *universes*.[19] A domain (as we use the term) is tied to the extension of a quantifier, relative to a specific context and interpretation (and any third parameter if such there be).[20] (Usually we derivatively attach the domain to the expression, the quantifier itself, relative to its interpretation and the context. When no confusion will arise we often leave the interpretation, and sometimes the context tacit). We have a fairly robust grasp of this notion. To give a circular elucidation: the domain of a quantifier in a context comprises the objects it ranges over (given its interpretation). For example, the domain of 'everything' in (2) relative to the context of utterance (in the relevant version of English) comprises the bottles and glasses, etc. on the table.[21]

A universe is tied to a whole language, the interpretation of an entire lexicon, *not*

[19] Although Fine does not gloss the distinction in terms of domains and universes, I think this provides a natural regimentation of the view he elucidates. Note that Fine uses 'absolute' non-generically as we use 'ultimate'. Setting aside the issue of Fine exegesis, this distinction will have an important role to play in the following.

[20] We shall usually leave any third parameter relativity tacit.

[21] A non-circular characterisation will follow in chapter 2.

relative to a specific context (but still relative to a third parameter if such there be). The universe of a language encompasses every object in the domain of any quantifier interpreted in the language in any context, together with any object that is the extension of a name or is a member of the extension of a predicate, and so on. The universe of the language in which (2) is uttered, the universe of a version of English, includes in addition to bottles and glasses, battleships and glaciers, and brittleness and gluons (assuming that there are such things) and whatever else we can refer to or quantify over in this language in any context. Consequently, relative to a fixed interpretation, the domain of a quantifier is always contained within the universe.

The distinction gives rise to a second decision point for the relativist. She must determine what prevents quantifiers achieving absolute generality. In order to succeed in quantifying over absolutely everything both an *intra*-language barrier and an *inter*-language barrier must be overcome. First the domain of the quantifier needs to be *unrestricted* over the universe of its interpretation: the domain must comprise every member of the universe. Second, the universe must be *inexpandable*: the universe must comprise every member of every universe. If a domain or universe comprises everything in every universe, it is open to neither further de-restriction nor expansion and I will accordingly label it *ultimate*. Quantifiers with an unrestricted (or ultimate) domain are said to be unrestrictedly (ultimately) general.

The absolutist maintains that both barriers may be overcome: some interpretations—including, he may add, the present version of English—have an inexpandable universe. Moreover in some contexts, no restrictions are imposed on the quantifier relative to this interpretation, and the domain is unrestricted. Unrestricted quantification over the inexpandable universe achieves ultimate generality.

Relativists maintain than one or other barrier precludes absolute generality. The *restrictionist*—or, to give her her full title, the anti-expansionist restrictionist—posits only the intra-language barrier. She claims that we cannot quantify over absolutely everything because the domain is never free from restriction. Even when—as she concedes is possible—the language is interpreted over the ultimate, inexpandable universe, the domain fails to be ultimate because it is a proper subdomain of this universe. The (anti-restrictionist) *expansionist* posits only the inter-language barrier. She claims that we cannot quantify over everything because the universe is always open to expansion. Even when—as she concedes is possible—the domain is unrestricted, comprising everything in the universe, the domain fails to be ultimate because the universe is a proper sub-universe of some other, more expansive universe.[22]

The expansionist's contention that even unrestricted quantifiers may be open to expansion may provoke an objection that it will be as well to dispense with immediately. The objection is that unrestricted quantifiers are trivially not open to expansion because any expansion would show the quantifier to have been *restricted* to a proper subdomain of the new universe after all. The objector might conclude that the contention that some quantifiers are both unrestricted and expandable is analytically false. It is no better than the dubious claim of an internet service provider that some of their packages are both unlimited and exhaustible (only that the latter has the good sense to hide the contradiction in the small print). Having the latter property trivially rules out having the former.

This objection betrays an important misunderstanding of the present use of 'unrestricted'. As we use the term, relative to a context C and an interpretation I, a quantifier as interpreted under I in C, is said to be *unrestricted* if the domain $D_{C,I}$ of the quantifier under I in C encompasses all of the universe M_I of the interpretation I. This is a special case of a more general notion. We can say that a quantifier as interpreted under I in C is *unrestricted over* M if its domain $D_{C,I}$ encompasses all of M. As such, to say that a quantifier is unrestricted *simpliciter* is simply to say that it is unrestricted over the universe of its interpretation. The equivocation in the objection is now clear. There is no contradiction in a quantifier's being both unrestricted *simpliciter*—unrestricted over its universe, M_I—and yet expandable, and so restricted relative to some other wider universe $M_{I'}$.

The absolutist may be tempted to rejoin that this is not how *he* understands 'unre-

[22]More extreme versions of relativism are also possible. The extreme relativist maintains that both barriers stand in the way of achieving absolute generality. For instance, it may be claimed both that domains are always restricted and that universes are always open to expansion. Even if we were *per impossible* to achieve unrestricted quantification, this would still fall short of quantification over the ultimate domain.

stricted'. On his view, it is an absolute notion, not tied to the present interpretation. To be truly unrestricted a quantifier must be unrestricted relative to the ultimate, inexpandable universe. But while, from the absolutist's point of view, such an absolute notion of 'unrestricted' is perfectly intelligible, this notion cannot figure in any dialectically effective argument against expansionism. For the intelligibility of the notion presupposes the existence of an inexpandable universe, and therefore presupposes the falsity of this view.

2.5 The varieties of relativism

Bringing both decision points together, relativism can be given a *provisional* statement as follows. (Stating relativism raises some difficulties that we shall return to in section 3.2 and at length in chapter 3.)

Relativism. No perspective's domain comprises absolutely everything.

This determines a family of views, generically labelled relativism about quantification. To determine a particular version of relativism, we must specify what accounts for shifts in perspective, and what prevents our achieving absolute generality. Several natural packages are available.

Contextual Restrictionism. One natural relativist package combines contextual and restrictionist relativism. Michael Glanzberg (2004, 2006) defends such a view: given any contextually determined domain, he takes it that a shift in context (leaving the interpretation fixed over the ultimate universe) can lead to a wider contextually determined domain. No domain is unrestricted in the ultimate universe.[23]

Procedural Postulationism. On the third-parameter view defended by Fine (2005, 2006, 2007), he rejects restrictionism and upholds expansionism: quantifiers can be unrestricted, with their domain encompassing the entire universe, but this universe itself is open to expansion by shifting the third parameter. We can always expand the universe by postulating further objects.

[23] We return to the details of this view in chapter 2.

Interpretational Expansionism. A third natural relativist package, and the view I shall defend, combines interpretational and expansionist relativism. Like Fine's view, I shall argue that although domains may be unrestricted, the universe is always open to expansion. But on my view what accounts for a shift in universe is a shift in interpretation rather than a shift in the seemingly mysterious third parameter. This view may also seem initially obscure. For without further elaboration, it is not much, if any, clearer how to effect reinterpretation-by-expansion than to carry out the non-circumstantial, non-contextual, non-interpretational shift apparently required by a procedural postulate. I think, however, that a clear account can be given. We shall return to this issue in chapter 6.

Sortal Restrictionism. A relativist moved by the sortal view espoused by P. T. Geach (1968) may sustain a distinctively interpretational version of restrictionism. It might be claimed, for instance, that since there is no universal sense of 'thing', the only way to confer meaning on quantifiers like 'everything' is to take 'thing' to have the same semantic value as some sortal predicate, like 'person' or 'cardinal number', that comes equipped with a non-trivial 'criterion of identity' for its instances. Shifts in perspective consist in shifts in the criterion of identity, by assigning 'thing' the semantic value of a different sortal predicate. Given that the extension of no sortal predicate includes that of every other, the domain of 'everything' so interpreted will never encompass every member of the universe. Assuming however that such shifts leave the overall stock of semantic values denoted by expressions in the lexicon unchanged, such a view counts as restrictionist since the universe remains the same.

Rather than focusing on the idiosyncrasies of particular packages we shall tend in part I to focus on the contextual/interpretation/third parameter and restrictionist/expansionist distinctions that underlie them. As we shall we see, these distinctions sometimes matter a great deal. But they will not always matter; and I shall continue to use the more generic terms 'relativism', 'perspective', and 'absolute' when distinctions between variants are unimportant.

3 Dialectical Issues

In the outline of relativism I have made free and uncritical use of quantification over domains and universes. This raises two dialectical issues that it would be as well to raise sooner rather than later.

3.1 The All-in-One Principle

The first is that how domains are construed is a delicate matter for the absolutist. For as is very widely appreciated,[24] the absolutist cannot in general treat domains as sets. Otherwise his characteristic claim that some domain contains everything is in conflict with standard, Zermelo-Fraenkel set theory, or its familiar extensions with large cardinal axioms, which refute the existence of a universal set.[25] Applying the Separation axiom, we obtain the Russell set of all sets that lack themselves as elements, which leads to a contradiction as above.

Similar trouble arises if we attempt to a construe a domain as any set-*like* object, a class or 2-class, or totality, multiplicity or what have you. Granted the analogue of Separation for domains—that whenever some things constitute a domain of a quantification, any things amongst them constitute another domain, a subdomain of the first, needed to allow for the domains of restricted quantifiers—the universal domain posited by the absolutist leads to the Russell domain, and the Russell domain leads to contradiction.

Richard Cartwright (1994, §§IV-V) has provided robust criticism of what he dubs the 'All-in-One Principle', the principle that 'to quantify over certain objects is to presuppose that those objects constitute a "collection," or a "completed collection"—some one thing of which those objects are the members' (p. 7) He writes:[26]

---

[24]See, e.g., Cartwright (1994, p. 7), Williamson (2003, p. 425), and Rayo and Uzquiano (2006, p. 6).

[25]Set theories such as NBG, or MK, that add classes to ZFC may countenance the existence of a class that has every set as an element. But such a class is not truly universal for it does not have proper classes such as itself as elements. To allow for a truly universal set, we would have to depart from mainstream set theory and move to a theory like NF which gives up on Separation.

[26]See also George Boolos (1993, pp. 223–4).

> Consider what it implies: that we cannot speak of the cookies in the jar unless they constitute a set; that we cannot speak of the natural numbers unless there is a set of which they are the members; that we cannot speak of all pure sets unless there is a class having them as members. I do not mean to imply that there is no set the members of which are the cookies in the jar, nor that the natural numbers do not constitute a set, nor even that there is no class comprising the pure sets. The point is rather that the needs of quantification are already served by there being simply the cookies in the jar, the natural numbers, the pure sets; no additional objects are required. (Cartwright 1994, p. 8)

Consequently, in the absence of independent support for this principle, relativist attacks on absolutism based on it will be dialectically ineffective and I shall not offer any such. All the same, we should be clear about just what this principle amounts to.

First, to reject the use of sets, or set-like objects, to encode domains is not to reject 'domains' altogether. For, as Cartwright concedes, what we do require in order to quantify over certain objects is for there to be *those objects*. Thus we can treat the objects, severally, rather than some one single set-like object, as encoding the domain of a quantifier. Rejecting the All-in-One Principle does not avoid the 'All-in-Many' principle that quantification over some things presupposes that there be those things.[27] The absolutist can make good sense of talk of 'domains' and 'universes' by rendering it in plural (or alternatively, second-order) terms.[28]

Second, and this is again a point freely admitted by Cartwright, to deny that quantification over some things *presupposes* that there be a set of those things is not to deny that there is a set of those things.[29] In particular, whenever some things are quantified over, the relativist may maintain that we can come to some more liberal domain which contains the set of those things—not because they were quantified over, but because on her view any things can thus form a set. One need not endorse the All-in-One Principle, to think that whatever can be quantified over can—from the perspective

[27]We borrow the label from Uzquiano (2009, p. 312). If the All-as-Many principle is not to collapse into the All-in-One principle, the absolutists needs to maintain, contrary Quine (1973, 1974), that first-order plural quantification over objects in English, is not disguised first-order singular quantification. See Boolos (1984, 1985) for a case in favour of treating plural quantification in plural terms.

[28]Gabriel Uzquiano (2009, p. 302) encodes domains plurally.

[29]See also Lavine (2006, p. 145).

of a wider domain—comprise a set-domain.[30]

On a terminological note, for ease of exposition, I shall continue to use singular quantification over and singular reference to domains and universes, and so on, leaving the absolutist to translate as we go along.

3.2 The liberal metalanguage assumption

A second issue, this time for the relativist, is that—by her lights—there is no perspective relative to which every member of any domain is quantified over. But given this, many of the relativist's attempts to deploy quantification in order to elucidate aspects of her view—including the previous sentence!—threaten to fall wide of the mark. The absolutist might state the problem as follows.

> By uttering 'there is no perspective from which every member of any domain is quantified over' the relativist wants to say that in no context—or interpretation, or whatever—can she quantify over everything. But this statement too is made in a context, $C^\star$, say. And by her lights this context's domain $D_{C^\star}$ falls short of encompassing everything. Instead, 'every member' quantifies only over members of $D_{C^\star}$, and 'every domain'—treating domains plurally—only over subdomains of $D_{C^\star}$. But then the utterance says that there is no context C—strictly, no *sub*-context C of $C^\star$—in which every member of $D_{C^\star}$ that is a member of any subdomain D of $D_{C^\star}$ is quantified over. And this is not only insufficiently general, but trivially false: for the context of utterance $C^\star$ itself is such a (sub-)context.

For expository purposes it will often be convenient therefore to suppose that the perspective from which I am writing—the version of English, context and so on, that then obtains—is more liberal than that of any of the perspectives—'object perspectives'—about which I am writing. (This remains convenient, even when as will often be the case in the following some of the perspectives I am writing about are also themselves 'meta-perspectives', perspectives relative to which one might theorise about other perspectives.) Assuming this is the case for the first sentence of this section, it does not say something trivially false, at the cost of quantifying slightly less generally over domains than is possible from this perspective when they are encoded plurally. The troublesome

[30]Compare Fine (2006, p. 24).

subsentence tells us only that there is no object perspective whose domain comprises every object in every other object perspective domain.

Similarly understood, the elucidation of the various relativist views set out in section 2.4 can be taken by their respective protagonists as correct *partial* descriptions of the structure of perspectives and domains, as they see them. And these may be supplemented by remarks—or even shifts in the meta-perspective—to illustrate that the described substructure is representative of the larger situation: analogous claims will be seen to hold of larger and larger potions of the hierarchy of domains carrying on up to and beyond any meta-perspective the relativist uses to describe parts of their view, and so on.

This will do for the present, provisional outline of these views. But the relativist should strive for more. So long as the partial characterisation of relativism fails to delimit the non-absoluteness of the domain of the most liberal perspective she ascends to in the course of setting our this characterisation, it leaves open the possibility that this very domain comprises absolutely everything. We shall return to this issue in chapter 3.

Part I

Generality

2 The Semantics of Quantification

The importance of the distinction between restrictionist and expansionist versions of relativism becomes very clear when we attend to the semantics of quantifiers. In the first place, the two views differ on the semantics itself: restrictionism claims all quantifiers are restricted; expansionism allows that quantifiers may be unrestricted. But, the rival formulations of relativism also lead to differences in how we can engage in semantic theorising, in what sorts of semantic theories are available to us.

On this second point, it has sometimes been argued that semantic theorising is one area where the absolutist secures a substantial advantage over the relativist: while the absolutist is well-equipped to theorise about the world as he sees it, the relativist struggles, given the limitations she posits, to delineate the semantic behaviour of quantifiers as they behave according to her theory. The following theses may be thought to hold.

Semantic optimism for absolutists. The absolutist *can* (by his lights) give adequate semantic theories for languages that deploy *absolutely* general quantifiers in a suitable metalanguage.

Semantic pessimism for relativists. The relativist *cannot* (by her lights) give adequate semantic theories for languages that deploy *non*-absolutely general quantifiers in a suitable metalanguage.

In particular, Williamson argues that the absolutist has no trouble in formulating the truth-conditions for sentences containing quantifiers with absolute domains in the familiar quasi-homophonic style associated with Tarski (1933) and Davidson (1967).

An absolutely general quantifier $\forall v$ can be given the following semantics.[1]

> $\forall v\phi$ is true under assignment σ iff everything a is such that ϕ is true under $\sigma[v/a]$

The absolutist must, of course, take 'everything' to likewise range over the absolute domain. (2003, p. 418).[2]

But, at least for contextual variants of relativism, this style of semantic theorising is problematic. Williamson argues as follows. The obvious Tarskian clause to allow for the contextual restrictions the relativist posits is the following one, generalising over contexts as well as assignments:

> $\forall v\phi$ is true in C under assignment σ iff every member a of the domain of C is such that ϕ is true under $\sigma[v/a]$

But, according to the contextual relativist, this clause must be set out in some theoretical context—$C^\star$, say—in which the metalanguage quantifier 'every member' ranges over a restricted domain, $D_{C^\star}$. And this will deliver the intended truth-conditions only if $D_{C^\star}$ is at least as inclusive as each object context's domain D_C. If not, $\forall v\phi$ may be true (under σ) in a context C on the basis of every member a of the *meta context's domain* in the domain of C—that is, every member of $D_{C^\star} \cap D_C$—being such that ϕ is true (under $\sigma[v/a]$) even though there is a counterinstance b in the domain of C such that ϕ is not true (under $\sigma[b/v]$). (2003, pp. 444-5)

This yields immediate trouble for the restrictionist. As Williamson observes, assuming the clause is adequate, $C^\star$ is unrestricted—contrary to restrictionism—provided only that there is no object that is outside each context's domain D_C (p. 445-6). But the (anti-expansionist) contextual restrictionist *does* contend this: she rejects the existence of what Williamson dubs 'semantic pariahs' and allows that some languages have an ultimate universe. Consequently, this view rules out giving an adequate semantics in this fashion.

[1]We adopt the logician's convention of omitting quotes from expression from (interpreted) formal languages. It will be clear from the context whether the expression is being mentioned, as in instances of the semantic clause below, or used. Quotes are still used to refer to English expressions.

[2]Assignments may be treated in the standard way as (set-)functions from variables to objects. As usual, the assignment $\sigma[v/a]$ agrees with σ on every variable other than v and maps v to a.

With respect to semantic optimism and pessimism, the absolutist has got off to a good start and the relativist to a bad one. But Tarskian semantic theories for restricted and unrestricted quantifiers stated in languages deploying such quantifiers are not representative of the wider situation. The semantic optimism and pessimism theses are sensitive both to the sort of semantic theory and to the sort of relativism that are in question. When we consider less Davidsonian approaches to semantics and expansionist approaches to relativism things look rather different.

The plan is this. The next section outlines and motivates two alternative approaches to the semantics of quantifiers in formal and natural languages: the categorial approach due to David Lewis (1970) and Richard Montague (1970), and the extensional approach associated with Mostowski (1957) and Lindström (1966), and famously applied to natural language by Barwise and Cooper (1981). With these approaches in hand, the second section compares the semantic commitments of restrictionism, expansionism, and absolutism. Finally, in the third section, we return to the issue of formulating semantic theories raised here to see how all sides fare when it comes to framing semantic theories in the extensional and categorial style.

1 Categorial and Extensional Semantic Theories

Compositional, truth-conditional semantics in the referential tradition of Frege (1892)—henceforth, simply: semantics—places two desiderata on a semantic theory.

Truth-conditionality. A semantic theory for a language must determine semantic values for each sentence ϕ of the language, which determine the correct truth-conditions for ϕ in each context of its utterance.

Compositionality. The semantic theory for the language must determine the semantic values of each complex expression ε as a function of the semantic values of the constituents of ε and their syntactic combination.[3]

A categorial approach to syntax and semantics provides a means to fulfil both desiderata, and a flexible framework within which to investigate the semantics of quantification.

[3]Lewis (1980, p. 25) formulates these desiderata in essentially this fashion.

1.1 Categorial Syntax

Let us start with a brief review of categorial syntax and semantics as presented by Lewis (1970).[4] Expressions are assigned to categories. There are a small number of *basic categories*. In fact, to start with, we shall deploy just two basic categories: the category of (open and closed) sentences, t, and the category of singular terms, e.[5] Derived categories are then generated as follows: whenever c and $c_1, \ldots, c_n$ are categories, then $c/(c_1, \ldots, c_n)$ is a derived category.

A categorial syntax specifies a lexicon consisting of *simple* expressions and assigns each a basic or derived category. Complex expressions are then formed and assigned categories according to a single formation rule:

Complex expressions. Whenever ε is an expression of category $c/(c_1, \ldots, c_n)$ and $\varepsilon_1, \ldots, \varepsilon_n$ are expressions of category $c_1, \ldots, c_n$, then $\varepsilon(\varepsilon_1, \ldots, \varepsilon_n)$ is an expression of type c. (1970, p. 20)

As usual, expressions which yield a sentence when supplied with n singular terms—that is, expressions of type $t/(e, \ldots, e)$ (with n occurrences of 'e', abbreviated t/e^n)—are called (n-ary, first-order) predicates. We often use p as short for t/e.

We shall focus on two languages containing quantifiers. The first is a prime candidate for a language that purports to achieve absolute generality, the language used to formulate Zermelo-Fraenkel set theory with urelements, $\mathcal{L}_{ß,\in}$.[6] The lexicon of this language is displayed in table 2.1. Other first-order languages are likewise given by adding constant symbols (category e), further or different relation symbols (category t/e^n), or function symbols (category e/e^n).

[4]Categorial grammar was first developed by Ajdukiewicz (1935) and Bar-Hillel (1964); aside from a few minor deviations in details, and changes in notation, we follow Lewis's presentation, adapting it to the examples that concern us. Page references in this section and the next are to his 1970 paper unless otherwise indicated.

[5]These are the labels deployed by Montague (1970): 't' stands for truth-value, 'e' for entity—viz. object, the extensions expressions of the categories receive. We can expect natural language to require more categories, but two will suffice for present purposes.

[6]We focus on set theory *with urelements* (non-sets), ZFCU, rather than the standard theory of *pure* sets, ZFC, as the latter does not purport to achieve absolute generality. Its quantifiers are implicitly restricted to range only over pure sets.

Table 2.1: The lexicon of $\mathcal{L}_{ß,\in}$

Category e	first-order variables: $x, y, z, x_1, y_1, z_1, \ldots$
Category t/e	unary predicate: ß (read: 'set') .
Category $t/(e,e)$	binary predicates: $=, \in$
Category $t/(t,t)$	binary connective: $\wedge$
Category t/t	unary connective: $\neg$
Category t/t	unary quantifiers: $\forall x$, $\forall y$, $\forall z$,. . .

In this case, and all others, all complex expressions are formed in accordance with the single syntax rule. For example, from the binary predicate $\in$ (category $t/(e,e)$) and the variable x (category e), the (open) sentence $\in(x,x)$ (category t) is formed. With this sentence and the quantifier $\forall x$ (category t/t), another sentence $\forall x(\in(x,x))$ is formed; and so on. To improve readability, we shall introduce the following abbreviations for expressions in the metalanguage: $\in(x,x)$ is written as $x \in x$; $\forall x(\in(x))$, as $\forall x(x \in x)$; and so on. Other connectives and quantifiers are introduced in the usual way, and further brackets are dropped in accordance with the usual conventions.

Unlike in $\mathcal{L}_{ß,\in}$, where quantifiers are variable-binding sentence operators, in English, quantifiers are noun phrases. A quantifier consists of a determiner ('every', 'some', 'no', 'most', etc.), together with a nominal, a singular or plural noun, perhaps qualified with adjectives or relative clauses ('thing', 'sailor', 'terrestrial donkey', 'bottles that are on the table', etc.). Quantifiers combine with verb phrases to form sentences.[7]

The lexicon of our second language approximates this syntax more closely. It is a categorial correlate of the formal languages used in Barwise and Cooper's seminal investigation of so-called 'generalised quantifiers' (1981, pp. 167–8), and we shall accordingly label it $\mathcal{L}_{\text{GQ}}$. Its core lexicon is displayed in table 2.2 (we shall add further predicates as required). In analogy with English, in $\mathcal{L}_{\text{GQ}}$, quantifiers (category $t/(t/e)$) are formed by combining a determiner (category $(t/(t/e))/(t/e)$) with a unary predicate (category t/e);[8] a quantifier together with a further unary predicate then yields a sentence.

Unlike $\mathcal{L}_{ß,\in}$, where variable binding is carried out by quantifiers, in the usual

[7]See, for instance, Lewis (1970, p. 40), Barwise and Cooper (1981, pp. 161–2).

[8]An alternative is to treat determiners as binary quantifiers, which in the present syntax fall into category $t/(t/e, t/e)$. But this leads to no substantive difference in semantics. See, for instance, Peters and Westerståhl (2006).

Table 2.2: The core lexicon of $\mathcal{L}_{\text{GQ}}$

Category e	first-order variables: $x, y, z, x_1, y_1, z_1, \ldots$
Category t/e	unary predicate: THING
Category $t/(t, t)$	binary connective: $\wedge$,
Category t/t	unary connective: $\neg$
Category $(t/e)/t$	abstraction operators: $\hat{x}$, $\hat{y}$, $\hat{z}$,$\ldots$
Category $(t/(t/e))/(t/e)$	determiners: EVERY, SOME, NO, MOST

way, in $\mathcal{L}_{\text{GQ}}$ it is achieved with abstraction operators: $\hat{v}$ combines with a sentence ϕ to form a unary predicate $\hat{v}(\phi)$, often written $\langle\hat{v}.\phi\rangle$, binding all free occurrences of v in ϕ. (Informally, $\hat{v}(\phi)$ can be read 'is such that $\phi(v)$' or 'satisfies $\phi(v)$'.)

This syntax provides both some syntactic economy,[9] and the means to make more fine-grained syntactic distinctions in the formalisation of quantified English sentences.[10] For instance, in a first-order language both 'Some bottle is empty' and 'Something is a bottle and empty' are formalised by applying a variable binding quantifier $\exists x$ to the open sentence $\textsc{bottle}(x) \wedge \textsc{empty}(x)$. In the new language, the syntactic differences between the two English sentences are brought out in two distinct formalisations:

$$\textsc{some}(\textsc{bottle})(\textsc{empty})$$

$$\textsc{some}(\textsc{thing})(\hat{x}(\textsc{bottle}(x) \wedge \textsc{empty}(x)))$$

with the variable binding operator $\hat{x}$ being used to form a complex predicate by binding the free variable in the open sentence $\textsc{bottle}(x) \wedge \textsc{empty}(x)$ only in the second case.

And in some cases this is not a mere economy: the abstraction strategy fails altogether.[11] Sentences like 'Most bottles are empty', cannot be adequately treated as the result of applying a syntactically unrestricted quantifier $\textsc{most}(\textsc{thing})$ to a predicate formed from abstraction on a sentence that combines $\textsc{bottle}(x)$ and $\textsc{empty}(x)$ using truth-functional connectives.[12]

[9] See Lewis (1970, p. 43).
[10] See Barwise and Cooper (1981, p. 165).
[11] See Lewis (1970, p. 43).
[12] Rescher (1962, p. 374) makes the observation that 'most' cannot be characterised with unary quantifiers and first-order resources; Kolatis and Väänänen (1995, thm. 5.5) prove a generalised version of this claim.

1.2 Categorial Semantics

Categorial semantics provides a direct means to meet the two desiderata on semantic theorising. Central to Lewis's categorial approach to semantics is the notion of an *index*. Fix a universe M. Let the extension of a singular term be a member of M and the extension of a sentence be its truth-value. Indices are n-tuples that encode features which may contribute to determining the extensions of singular terms and sentences. Since these extensions are sensitive to how the world is, an index i must incorporate a possible world co-ordinate w_i. As we shall see below, for the languages that interest us here, it is also necessary for indices to incorporate an assignment co-ordinate σ_i, which maps each variable to a member of M.

In order to meet the first desiderata, we need to determine the truth-conditions for each sentence in each context. Accordingly, in each context C, the semantic value of a sentences ϕ, is a function $[\![\phi]\!]_C$ that maps indices to truth-values.[13]

The first desideratum is therefore met. Similarly the semantic value of other expressions from basic categories in each context is taken to be a function from indices to suitable extensions. Lewis labels such functions *Carnapian intensions.* Semantic values are divided into the same categories as expressions. For example, a Carnapian intension from indices to truth-values, the semantic value of a sentence, a category t expression, is a category t semantic value. We use boldface symbols $\boldsymbol{t}$, $\boldsymbol{e}$ for category t and e semantic values. To summarise so far then we have:

Intensions for basic categories. A category e intension is a function $\boldsymbol{e}$ that maps each index to a member of the object language universe M. A category t intension is a function $\boldsymbol{t}$ that maps each index to a truth-value, T or F.[14]

The semantic values of derived expressions are picked with the second desideratum in mind. This requires that the semantic value of a complex expressions be determined

[13]This is a departure from Lewis's (1970) presentation in which features of context, such as speaker, are included as co-ordinates in indices. However in his 1980 paper, Lewis advances the view that there is no deep difference between a theory that assigns different semantic values to expressions in different contexts, as we do here, and one that incorporates a context parameter C as a further co-ordinate in indices.

[14]See Lewis (1970, §III)

Table 2.3: Intensions for simple $\mathcal{L}_{\beta,\in}$-expressions

Write $i[v/a]$ for the index that is the same as i save that its assignment is $\sigma_i[v/a]$ and let M_{w_i} comprise the members of M that exist at w_i.

$$\begin{aligned}
[\![v]\!](i) &= \sigma_i(v) \\
\big([\![=]\!](\boldsymbol{e}, \boldsymbol{e}')\big)(i) &= T \text{ iff } \boldsymbol{e}(i) = \boldsymbol{e}'(i) \\
\big([\![\in]\!](\boldsymbol{e}, \boldsymbol{e}')\big)(i) &= T \text{ iff } \boldsymbol{e}(i) \in \boldsymbol{e}'(i) \\
\big([\![\neg]\!](\boldsymbol{t})\big)(i) &= T \text{ iff } \boldsymbol{t}(i) = F \\
\big([\![\wedge]\!](\boldsymbol{t}, \boldsymbol{t}')\big)(i) &= T \text{ iff } \boldsymbol{t}(i) = T \text{ and } \boldsymbol{t}'(i) = T \\
\big([\![\forall v]\!](\boldsymbol{t})\big)(i) &= T \text{ iff } \boldsymbol{t}(i[v/a]) = T \text{ for every } a \in M_{w_i}
\end{aligned}$$

from the semantic values of its constituents. For instance, compositionality demands that we should determine the semantic value of a sentence (category t) of the form $F\tau$ on the basis of the semantic value of F (category t/e) and the semantic value of τ (category e). The categorial approach to semantics achieves this by taking category t/e semantic values, to be functions that map category e semantic values to category t semantic values. Lewis labels such functions compositional intensions. More generally, the compositional intensions for expressions from derived categories are as follows.

Intensions for derived categories. A category $d = c/(c_1, \ldots, c_n)$ intension is a function $\boldsymbol{d}$ that maps a category c_1-intension $\boldsymbol{c_1}$, …, c_n-intension $\boldsymbol{c_n}$ to a c-intension $\boldsymbol{c}$.[15]

In order to give a categorial semantics for a particular lexicon we need to specify a category c semantic value for each simple category c expression in the lexicon. For example, intensions for the simple expressions of $\mathcal{L}_{\beta,\in}$ and $\mathcal{L}_{\mathrm{GQ}}$ are specified in tables 2.3 and 2.4. (For now we ignore the effects of context on quantifiers; we shall return to this issue in section 2.)[16] The semantic values of complex expressions are provided by a single compositionality clause.

Compositionality. Whenever $\boldsymbol{d}$ is a category $d = c/(c_1, \ldots, c_n)$ intension and $\boldsymbol{c}_1$,… $\boldsymbol{c}_n$ are intensions from categories $c_1, \ldots c_n$, the intension that results from composing

[15] See Lewis (1970, §IV).

[16] The semantics we give here for SOME and the abstraction operator are in essentials the same as those given by Lewis (1970, pp. 41, 45), making minor adaptations for the language, and rather more substantial changes to the notation. The semantics for $\forall$ is approximately the dual of the semantics Lewis gives for "the logician's 'some"' (p. 42).

these intensions is the category c intension—written: $\boldsymbol{d}\langle\!\langle \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \rangle\!\rangle$—obtained by function-application: $\boldsymbol{d}\langle\!\langle \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \rangle\!\rangle = \boldsymbol{d}(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)$.

This provides an equally direct means to meet the second desideratum. The intension of a complex expression $\varepsilon(\varepsilon_1, \ldots, \varepsilon_n)$ is taken to be the result of composing the intensions of its constituents: $[\![\varepsilon(\varepsilon_1, \ldots, \varepsilon_n)]\!] = [\![\varepsilon]\!]\langle\!\langle [\![\varepsilon_1]\!], \ldots, [\![\varepsilon_n]\!] \rangle\!\rangle$ [17]

Two comments are in order. Notice first that we will only meet both desiderata if the indices are rich enough. Consider for instance $\forall x ß(x)$. Under the specified semantics we recover the intended intension $[\![\forall x ß(x)]\!]$—a function that maps i to truth just in case every member of the universe M_{w_i} of w_i is a set at w_i—as a function of the intensions of its immediate constituents, $[\![\forall x]\!]$ and $[\![ß(x)]\!]$. Even though the truth-value of the final sentence does not depend on the assignment, the inclusion of an assignment co-ordinate in the index is crucial to recovering the correct truth-conditions compositionally. For the truth-value of the whole sentence at an index i depends on the truth-value of the subsentence $ß(x)$ not just at the actual index, but at every index $i[a/x]$ that differs from i in the assignment it makes to x. If indices did not incorporate an assignment co-ordinate, and the intension of the subsentence $[\![ß(x)]\!]$ did not encode its truth-value at each assignment, there would be no function to play the role of $[\![\forall x]\!]$, determining the intended intension of the whole sentence from the intension of the subsentence. For a less powerful language fewer indices may do. For instance, the semantics of the language of propositional logic would not require an assignment co-ordinate to meet the desiderata. On the other hand, more powerful languages will require richer indices. The introduction of tense operators for instance would require us to include a time co-ordinate. For our purposes here it will suffice to take indices to be world-assignment pairs $\langle w_i, \sigma_i \rangle$.[18]

Notice second that we rely on functions that are built over the initial universe M. Lewis is quite explicit that M should be a set, and the functions defined set-theoretically as usual. (See (1970, pp. 23, 26).) But such an approach will not be acceptable to the absolutist when he takes the object language to be interpreted over the absolute universe. For no set has everything as an element. We shall return to this issue in section 3.

[17] See Lewis (1970, §IV).

[18] Lewis (1970, p. 25) tentatively suggest eight indices for natural language.

Table 2.4: intensions for $\mathcal{L}_{\text{GQ}}$ determiners and $\hat{v}$

Write $a \prec_{\text{i}} \boldsymbol{p}$ to abbreviate $\exists \boldsymbol{e}(a = \boldsymbol{e}(i) \wedge (\boldsymbol{p}(\boldsymbol{e}))(i) = T)$; and let M_{w_i} comprise the members of M that exist at w_i.

$$\big(\llbracket\text{EVERY}\rrbracket(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff every member } a \text{ of } M_{w_i} \text{ s.t. } a \prec_{\text{i}} \boldsymbol{p} \text{ is s.t. } a \prec_{\text{i}} \boldsymbol{p}'$$
$$\big(\llbracket\text{SOME}\rrbracket(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff some member } a \text{ of } M_{w_i} \text{ s.t. } a \prec_{\text{i}} \boldsymbol{p} \text{ is s.t. } a \prec_{\text{i}} \boldsymbol{p}'$$
$$\big(\llbracket\text{NO}\rrbracket(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff no member } a \text{ of } M_{w_i} \text{ s.t. } a \prec_{\text{i}} \boldsymbol{p} \text{ is s.t. } a \prec_{\text{i}} \boldsymbol{p}'$$
$$\big(\llbracket\text{MOST}\rrbracket(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff most members } a \text{ of } M_{w_i} \text{ s.t. } a \prec_{\text{i}} \boldsymbol{p} \text{ are s.t. } a \prec_{\text{i}} \boldsymbol{p}'$$
$$\big((\llbracket\hat{v}\rrbracket(\boldsymbol{t}))(\boldsymbol{e})\big)(i) = \boldsymbol{t}(i[\boldsymbol{e}(i)/v])$$

1.3 Extensional semantics

The categorial semantics for $\mathcal{L}_{\text{GQ}}$ can be replaced with a semantic theory of a more familiar kind for this language. Barwise and Cooper (1981) follow a largely extensional approach to giving a semantics for $\mathcal{L}_{\text{GQ}}$.[19]

Taking it for granted that there is a set comprising the members of the object language universe, this approach treats the extensions of singular terms and sentences as objects and truth-values, as in the categorial approach, but also assigns extensions to predicates and quantifiers. Each n-ary predicate θ is given an n-ary relation $|\theta|$ over the set-universe as its extension, comprising the extensions of terms to which the predicate applies; each (unary) quantifier Q is given a set of subsets of the universe $|\text{Q}|$ as its extension, comprising the extensions of predicates to which the quantifier applies; and each determiner D is assigned a function $|\text{D}|$ as its extension that maps predicate-extensions to quantifier-extensions. For example, the extension of the quantifiers EVERY(THING) and SOME(THING) are the sets: $\{M\}$ and $\{A \subseteq M \mid A \neq \varnothing\}$. The truth-values of sentences are then determined compositionally from the extensions of their constituents, so that for instance:

$$\begin{aligned} |\text{EVERY(THING)(BOTTLE)}| = T &\text{ iff } |\text{BOTTLE}| \in |\text{EVERY(THING)}| \\ &\text{ iff } |\text{BOTTLE}| = M \\ &\text{ iff every element of } M \text{ is in } |\text{BOTTLE}| \end{aligned}$$

[19]Page references in this section are to Barwise and Cooper's 1981 paper

Table 2.5: Extensions of $\mathcal{L}_{\text{GQ}}$ determiners

$$
\begin{aligned}
|\textsc{every}|_i(A) &= \{B \subseteq M_{w_i} \mid A \subseteq B\} \\
|\textsc{some}|_i(A) &= \{B \subseteq M_{w_i} \mid A \cap B \neq \varnothing\} \\
|\textsc{no}|_i(A) &= \{B \subseteq M_{w_i} \mid A \cap B = \varnothing\} \\
|\textsc{most}|_i(A) &= \{B \subseteq M_{w_i} \mid |A \cap B| > |A \backslash B|\}
\end{aligned}
$$

To avoid presupposing a set-universe, let us generalise their approach. Provisionally we replace extensions of predicates with classes of objects, or 1-classes, and extensions of quantifiers with classes of classes of objects, 2-classes, with the expectation that the absolutist will ultimately want to eliminate class talk much as he eliminates universe talk.

Slightly generalising Barwise and Cooper's semantics in this way, we are able to replace the more complicated intensions of the terms, predicates and determiners of $\mathcal{L}_{\text{GQ}}$ with simpler extensions, varying across different indices, and contexts, at the cost of increasing the number of clauses governing compositionality. The extensions for determiners—still ignoring contextual domain restriction—are displayed in table 2.5.[20] And in addition to the usual Tarskian clauses for connectives, the extension $|\phi|_i$ of each sentence ϕ is determined as follows. (Compare Barwise and Cooper (1981, §2.5).)

Extensions for sentences.

C1 $|\theta(\tau_1, \ldots, \tau_p)|_i = T$ iff $\langle |\tau_1|_i, \ldots, |\tau_p|_i \rangle \in |\theta|_i$

C2 $|\textsc{d}(\theta)(\theta')|_i = T$ iff $|\theta'|_i \in |\textsc{d}|_i(|\theta|_i)$

A pleasing feature of the categorial approach to the semantics of determiners is that the extensional approach follows as a special case whenever it is feasible. Say that a predicate intension $\boldsymbol{p}$ (category t/e^n) is *extensional* if for each index i there is a class $|\boldsymbol{p}|_i$ such that:[21]

$$\big(\boldsymbol{p}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)\big)(i) = T \text{ iff } \langle \boldsymbol{e}_1(i), \ldots \boldsymbol{e}_n(i) \rangle \in |\boldsymbol{p}|_i.$$

[20] The extension for MOST is for the weakest sense of 'most' as meaning more than half.

[21] Lewis (1970, pp. 29–30) gives a similar definition of extensionality for verb phrases.

Likewise, say that a determiner intension $\boldsymbol{d}$ (category $(t/(t/e))/(t/e)$) is *extensional* if, for each index i there is a function $|\boldsymbol{d}|_i$ such that, for any extensional predicate denotations $\boldsymbol{p}$ and $\boldsymbol{p}'$:

$$\big(\boldsymbol{d}(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff } |\boldsymbol{p}'|_i \in |\boldsymbol{d}|_i(|\boldsymbol{p}|_i).$$

So defined, $|\boldsymbol{p}|_i$ and $|\boldsymbol{d}|_i$ are unique if they exist. Consequently, we may refer to $|\boldsymbol{p}|_i$ as the extension of $\boldsymbol{p}$. Extensionality and extensions are derivatively attached to predicates relative to a categorial interpretation in the obvious way: θ is extensional if $[\![\theta]\!]$ is extensional; and if so, the extension of θ is $|[\![\theta]\!]|_i$, which we may write simply as $|\theta|_i$. Similarly for determiners.

Given these definitions, it is readily seen that the intensional semantics for the determiners of $\mathcal{L}_{\text{GQ}}$ yields the extensional semantics. The categorial interpretation of the determiners in table 2.4 ensures that these expressions receive the intended extensions over the universe, displayed in table 2.5; and the compositionality clauses for extensional expressions, (C1) and (C2), follow from the single categorial compositionality clause (on page 30).

But not every $\mathcal{L}_{\text{GQ}}$ expression is open to an extensional treatment. We have to make a special case of the variable binding operator $\hat{v}$. Extending our earlier definitions in the natural way, we cannot assign the variable binding operator $\hat{v}$ an extension because there is no function which determines the extension of $\langle \hat{v}.\phi \rangle$ on the basis of the extension (i.e. truth-value) of ϕ. Instead of assigning it an extension, we must give a separate semantic clause for $\hat{v}$, which determines the extension of the predicate $\langle \hat{v}.\phi \rangle$ at a given index i on the basis of the extension of ϕ at each index $i[v/a]$.[22]

Extensions for abstracted predicates. $|\langle \hat{v}.\phi \rangle|_i = \{a \mid |\phi|_{i[v/a]} = T\}$

For this reason extensions cannot be semantic values for $\mathcal{L}_{\text{GQ}}$-expressions in the sense specified in the second desideratum on semantic theories. The extension of an abstracted predicate is not a function of the extensions of its constituents. Still, for many purposes,

[22] Barwise and Cooper (1981) are not explicit about the semantics of the abstraction operator in the language; but see, for instance, Stalnaker (1977, p. 329).

we can focus on the extensions of determiners and quantifiers rather than their intensions.

2 Quantifier Domain Restriction

The categorial and extensional semantics for the quantifiers in $\mathcal{L}_{ß,\in}$ and $\mathcal{L}_{\mathrm{GQ}}$ presented so far will not be acceptable to all sides in the absolutism/relativism debate. One issue, mentioned already, concerns the metatheoretic resources used to frame the semantic theory, the use of functions and classes. But the restrictionist will object not just to the metatheory but also to the theory itself. Amongst other things, we have so far ignored context sensitivity, which according to one of the most promising variants of this view, plays a crucial role in excluding unrestricted quantification. In this section, we shall take up the issue of the semantics of quantifier domain restriction, returning to metatheoretic issues in the next.

2.1 Domains

Let us start with the notion of *domain*. Pre-theoretically, many quantifiers seem to be restricted to range only over a limited domain. 'Most bottles on the table are empty' seems in some sense to be restricted just to bottles on the table. But the extensional semantics for 'most' tells us that whether this sentence is true depends on whether a certain relation holds between the extension of 'bottle on the table' and the extension of 'empty'. Now the first extension falls within the intuitive domain; but the second extension does not. In what sense then is the quantifier restricted to bottles on the table?

The answer is that we do not need to look at the entire extension of 'empty', examining every empty thing in the entire world, in order to determine whether the relation holds. We need only look at that portion of the extension within the domain. Most bottles on the table are empty just in case most bottles on the table are empty bottles on the table. The relation characteristic of 'most' holds between the extension of 'bottle on the table' and the extension of 'empty' just in case it holds between the extension of 'bottle on the table' and the intersection of the extension of 'empty' with the domain.[23]

The same is true of a wide range of natural language determiners. Barwise and Cooper (1981) propose the following as a linguistic universal: when a natural language quantifier is formed from a determiner $\textsc{d}$ and a nominal predicate θ, the quantifier $\textsc{d}(\theta)$ lives on the extension of the nominal, which is to say that for any predicate extension B over the universe:[24]

$$B \in |\textsc{d}(\theta)| \text{ iff } B \cap |\theta| \in |\textsc{d}(\theta)|.$$

Consequently, taking the *domain* of a quantifier $\textsc{d}(\theta)$ (at index i and context C) to be the smallest predicate extension that the quantifier lives on, provides a natural regimentation of our pre-theoretic notion. In particular, for the determiners in $\mathcal{L}_{\text{GQ}}$, the extension of θ is the domain of the quantifiers $\textsc{every}(\theta)$, $\textsc{some}(\theta)$, $\textsc{no}(\theta)$, and $\textsc{most}(\theta)$.[25]

With this regimentation of the notion of domain, we likewise attain a regimentation of the characteristic thesis of restrictionism that no quantifier is unrestricted. The claim that there is no context such that the universal quantifier's domain encompasses the entire universe amounts to the claim that its extension always lives on a proper subdomain of the universe.[26] The absolutist, by contrast, will claim that, in some contexts, quantifiers like $\textsc{every}(\textsc{thing})$ do not live on any extension other than that of the universal predicate, $\textsc{thing}$, which encompasses the entire universe of the language. The restrictionist and absolutist thus disagree over a substantive thesis about the semantics of quantification; one which we might expect to be settled on semantic grounds.

2.2 Nominal Restrictions

Since the domain of $\textsc{every}(\textsc{thing})$, when interpreted as in the previous section to receive the extension $\{M\}$ *does* encompass the entire universe, the restrictionist will need to modify the semantic account given so far. The next three subsections will assess the

[23] Compare Williamson (2006, p. 449).

[24] This property is now known as conservativity. See Peters and Westerståhl (2006, §4.5) for an overview.

[25] While this notion of domain is well-defined in a wide range of cases, including all those that concern us here, it is possible to construct quantifiers such that there is no smallest predicate extension that they live on. See Peters and Westerståhl (2006, §3.2.2).

[26] We continue to defer the question of how the relativist is to state their view in good conscience until the next chapter.

salient options for the restrictionist to pursue.

The natural move for the *sortal* restrictionist to make is to contend that—contrary to the absolutist's claim about THING—there are no universal predicates whose extension encompasses the entire universe. The quantifiers in $\mathcal{L}_{\text{GQ}}$ always live on the extension of some sortal predicate θ, whose extension is a proper subdomain of the universe, not containing satisfiers of other sortal predicates.

Such a move however threatens to be unstable. It becomes unclear why we cannot construct a universal predicate, playing the role of THING that the sortal restrictionist sought to banish from the language, and on this basis recover an unrestricted quantifier. In any natural language there will only be finitely many sortal predicates: $\theta_1, \ldots, \theta_n$. But then we can form a universal predicate using the abstraction operator and disjunction.

$$\langle \hat{x}.\theta_1(x) \vee \ldots \vee \theta_n(x) \rangle$$

Supposing that every object falls into some sort, belonging to the extension of some θ_p, and granted the intensions set out for $\hat{v}$, $\vee$ and EVERY, are as tables 2.3 and 2.4, the quantifier formed from composing EVERY with this disjunctive predicate has the entire universe as its domain.

Assuming that the sortal restrictionist does not take issue with the semantics of disjunction or abstraction, two potential responses are apparent. The first is to impose a syntactic ban on complex predicates like the troublesome one composing with determiners. The second is to seek to impose further semantic restrictions on the determiner, in addition to those already imposed on the nominal predicate.

The first options seems to be a non-starter. Given that the determiner EVERY and the disjunctive universal predicate receive their intended intensions, there is nothing to stop these *intensions* being composed: applying the function that is the intension of EVERY to the intention of the predicate. And without a semantic reason to stop these semantic values being composed any syntactic ban appears hopelessly ad hoc.

The second options pushes the sortal restrictionist towards a contextual view. Without positing context sensitivity, the only intension that delivers the correct truth-conditions is the one deployed by the absolutist. To assign EVERY an intension with a further built in restriction, such as

$$\big(\llbracket \text{EVERY} \rrbracket^{D}(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff every member } a \text{ of } D \text{ s.t. } a \prec_{\mathrm{i}} \boldsymbol{p} \text{ is s.t. } a \prec_{\mathrm{i}} \boldsymbol{p}'$$

where D is a proper subdomain of M_{w_i}, will incorrectly rule EVERY$(\theta_1)(\theta)$ to be true when every satisfier of θ_1 *in* D satisfies θ despite there being a counterinstance that satisfies the nominal θ_1 but not θ.[27]

Plausible options for a purely sortal view seem thin on the ground. But the sortal restrictionist might move to a mixed view: rather than taking D to be fixed, the restrictionist might allow that it varies according to which nominal θ_1 is composed with the determiner.

2.3 Indexical Restrictions

Consider contextual restrictionism more generally. In order to maintain that contextual factors always limit the domain of our quantifiers to a proper subdomain of the universe, the most straightforward proposal for the restrictionist to make is that the determiner EVERY functions effectively as an indexical expression: it is restricted to ranging over a background domain D_C,[28] a proper subdomain of the universe M, in virtue of receiving different intensions in different contexts:

$$\big(\llbracket \text{EVERY} \rrbracket_{C}^{D_C}(\boldsymbol{p})(\boldsymbol{p}')\big)(i) = T \text{ iff every member } a \text{ of } D_C \text{ s.t. } a \prec_{\mathrm{i}} \boldsymbol{p} \text{ is s.t. } a \prec_{\mathrm{i}} \boldsymbol{p}'$$

While this provides a straightforward treatment of sentences like the examples in the introduction, there is agreement on both sides of the absolutism/relativism debate that this cannot account for the phenomena in general.[29]

Consider the following example from Stanley and Szabó (2000, p. 249). Suppose

[27] Compare Williamson (2003, §VII).
[28] We take this terminology from Glanzberg (2006, p. 49). Unlike Glanzberg, however, who uses M for contextually determined background domains—speaking of 'M-dependence' for sensitivity to this domain—we continue to use D for domains and to reserve M for universes.
[29] See Williamson (2003, p. 418–9) and Glanzberg (2006, pp. 51–2).

that as the boat leaves the harbour, all the sailors stand on deck and wave to all the sailors on the shore who wave back. It seems that in such a context, an utterance of:

(1) Every sailor waved to every sailor

says that every sailor *on the boat* waved to every sailor *on the shore*.[30] Or again, suppose that everyone who came along to the party brought some bottles with them, and for some reason or other drank only what they had contributed. It seems that uttered in such a context:

(2) Everyone drank every bottle

says that everyone *who came* drank every bottle *that they brought with them*.[31]

Neither of these readings can be recovered if we take contextual restrictions to determine the same domain for each quantifier in the sentence. For in a context C, such an account must take (1) to say that every sailor in D_C waved to every sailor in D_C; and (2) to say that every person in D_C drank from every bottle in D_C. Neither of these captures the intended reading for any choice of domain D_C. The former wrongly implies that the sailors in D_C waved to themselves; and the latter wrongly implies that everyone drank from the same bottles. (If in either case there are no sailors or people in the domain D_C we miss the intended reading by making the respective sentences vacuously true regardless of what the sailors or people did.)

2.4 Parameter Restrictions

The examples in the last section show that it is not enough for the restrictionist to posit indexicality in determiners. A second option is for her to maintain that rather than contributing to determining the intension of a constituent already explicit in the syntax, context contributes *further* constituents to semantic structure.

To obtain the correct readings in cases like these, Stanley and Szabó (2000, pp. 245–50) argue, different domains need to be associated with different quantifier occurrences;

[30] See also the example attributed to Peter Ludlow by Stanley and Williamson (1995).

[31] Williamson (2003, p. 419) presents a similar example; cf. the example presented by Stanley and Szabó (2000, p. 243).

in the case of (2), different domains need to be associated with the quantifier 'every bottle' for different people in the first-domain. We need the contextual restriction on 'every bottle' to be bound by the outer quantifier. The moral they draw from such examples is that contextual restrictions contribute constituents to semantic structure.

Stanley and Szabó's preferred proposal is that the context provides parameters that compose with nominals to restrict quantifiers.[32] (See pp. 251–3.) Applying this proposal to the second example, (2) receives a logical form corresponding to the $\mathcal{L}_{\text{GQ}}$ sentence:[33]

$$\textsc{every}\big(\textsc{one}, f(y)\big)\big(\langle\hat{x}.\textsc{every}(\textsc{bottle}, g(x))(\textsc{drank}(x))\rangle\big)$$

where the nominal predicates ONE and BOTTLE combine with further constituents $f(y)$ and $g(x)$ to form category t/e expressions $\textsc{one}, f(y)$ and $\textsc{bottle}, g(x)$ and the variable in $g(x)$ is bound by the abstraction operator $\hat{x}$. Provided the context determines the intensions of f and g so that the extension of $\textsc{one}, f(y)$ in i comprises the people who attend the party in w_i and the extension of $\textsc{bottle}, g(x)$ in i comprises the bottles brought by $\sigma_i(x)$ in w_i, we obtain the correct truth-conditions for (2).

While proposals along these lines avoid the semantic problems faced by the indexical account, and are therefore an appealing option for the absolutist to adopt, the restrictionist who attempts to implement this option faces a philosophical difficulty. The problem is similar to the one that faced the sortal restrictionist. If contextual domain restriction works solely by somehow composing a contextually supplied constituent with a contextually invariant determiner intension and a nominal intension, it is unclear why we cannot achieve unrestricted quantification simply by operating in a context in which either no further constituent is supplied or the constituent's extension encompasses the entire universe.[34] For, as in the sortal case, the only constant intension for EVERY that will succeed in delivering the correct truth-conditions in every context is the one deployed

[32] An alternative proposed by Westerståhl (1985a) is to take the further parameter to compose with the determiner. For present purposes the differences between these accounts are unimportant.
[33] To minimise on the need for abstraction we suppose that DRANK has category $(t/e)/e$.
[34] Compare Williamson (2003, p. 420).

by the absolutist. But once this much has been conceded, any bar on unrestricted quantification seems ad hoc. On this view, the unrestricted intension the absolutist attributes to EVERY and the universal one he attributes to THING are both available. The result of composing them therefore expresses unrestricted universal quantification over the entire universe. Why then are we unable to pick out this intension? The restrictionist is left to explain why, no matter what we say or do or intend, the context always slips in a further non-trivial constituent; and why we cannot attach this intension to quantifiers in artificial language such as $\forall v$ that are not subject to contextual restrictions.

2.5 Semantic Commitments

Restrictionist accounts of quantifier domain restriction are therefore bound by their philosophical commitments as well as the semantic data. In the first place, as we have presupposed in this section, quantifier domain restriction must be treated as a semantic phenomenon. Context must make a difference to what is said, contrary for instance to the purely pragmatic view defended by Bach (2000). So far this is perhaps no great cost.[35] But when choosing among different semantic accounts, the restrictionist is again limited by her philosophical commitments. A purely sortal account seems unworkable. A purely indexical view runs contrary to the data; and a purely parameter-based view gives up too much to the absolutist. It seems therefore, that the restrictionist must follow Glanzberg (2006) in maintaining that *both* determiner indexicality and contextual-parameters feature in quantifier domain restriction.[36] The conjunctive view thereby incurs substantial semantic commitments—notably, making the striking claim that every determiner is indexical—and loses out on simplicity in comparison to the absolutist who can claim that the phenomenon of quantifier domain restriction is wholly down to the composition of contextually supplied parameters with the fixed semantic values of a determiner and a nominal.

But a point that should also be emphasised at this stage is that these commit-

[35] The semantic view enjoys independent motivation. See Stanley and Szabó (2000).

[36] Glanzberg's reasons for positing indexical restrictions are differently motivated, due to background domains associated with the determiner in the context being large and largely unchanging compared with the local domains associated with particular quantifiers (2006, p. 51–4).

ments, and any attendant costs, are peculiar to *restrictionist* variants of relativism. The expansionist relativist agrees with the absolutist that quantifiers are sometimes unrestricted. She can likewise follow him in maintaining that quantifier domain restriction is purely due to contextual parameters, or even follow Bach's heroic stance. The restrictionist's disagreement with the absolutist is at bottom a disagreement about semantics—in Barwise and Cooper's terminology, it amounts to disagreeing about whether quantifiers always live on a proper subdomain of the universe. In contrast, the expansionist sides on this semantic point with the absolutist. The expansionist and absolutist disagree about *meta*semantics.[37] Where she differs from the absolutist is not over the intrinsic features of any given interpretation of the lexicon but in its relation to *other* interpretations. In particular, what the expansionist contends, and the absolutist denies, is that there is always a more inclusive interpretation with a wider universe. But while absolutists and relativists thus agree about the content of semantic theories they will disagree about how they should be formulated. This brings us to the final subject of this chapter.

3 Stating Semantic Theories

In stating the standard categorial semantics for the language of first-order set theory, $\mathcal{L}_{\beta,\in}$, we have made free use of functions. The intensions of expressions in basic categories are functions mapping indices to objects; the intensions of expressions in derived categories are functions mapping functions to functions. How are these functions to be understood?

The answer depends on the position adopted in the absolutism/relativism debate. The canonical approach, that categorial and extensional semantics are framed in set theory, is available in all cases only to the expansionist.

[37]Or perhaps on the sort of view defended by Fine (2005, 2006), according to which shifts in universe are due to shifts in the third parameter, the ontology, they disagree about metaphysics. See chapter 1, section 2.3.

3.1 Expansionist semantics

Focus on intended interpretations of the lexicon of $\mathcal{L}_{\beta,\in}$. Following Zermelo (1930), it is open to the expansionist to claim that there is no single intended interpretation but rather an unbounded sequence of more and more liberal set-theoretic universes. In particular, she can claim that for any given interpretation, we can come to another such that (i) the members of the object language universe form a set M—some inaccessible rank $V_\theta[U]$—in the new universe; (ii) the new universe is closed under set-theoretic operations, rendering the axioms of ZFCU true under the new interpretation; and (iii) the new interpretation stably expands the old one: elements of M satisfy the atomic predicates of $\mathcal{L}_{\beta,\in}$ in the new interpretation just in case they satisfy them in the old one. Taking such a language as the metalanguage provides a congenial setting for formulating a categorial semantic theory for the given interpretation in the canonical way within set theory.[38]

It is straightforward if tedious to encode categorial syntax in set theory. Without going into the details let us assume that we have encoded the syntax of $\mathcal{L}_{\beta,\in}$ in set theory so that we have a set of categories CAT and a set of expressions EXP, the second of which is a superset of both the set of variables VAR and the set of sentences SENT. Assume moreover that we have well-founded relations of derivedness and complexity defined in the natural way, so that $d = c/c_1, \ldots, c_n$ is more derived than its components: $c, c_1, \ldots, c_n$, and $\varepsilon(\varepsilon_1, \ldots, \varepsilon_n)$ more complex than its constituents: $\varepsilon, \varepsilon_1, \ldots, \varepsilon_n$. With the syntax in hand, turn to the semantics.

As usual, the requisite intensions are encoded as set-functions. But the familiarity of this approach should not lead us to forget that a fair amount of *coding* is deployed.

Encoding functions as sets. First we define the Kuratowski ordered pair of x and y in the usual way: $\langle x, y\rangle =_{\mathrm{df}} \{\{x\}, \{x, y\}\}$. This definition permits us to prove the pairing theorem for Kuratowski pairs:

$$\langle x_1, y_1\rangle = \langle x_2, y_2\rangle \text{ iff } x_1 = x_2 \text{ and } y_1 = y_2.$$

[38]Montague (1970) provides an account of how to encode categorial syntax and semantics in set theory, similar in essentials to the one presented in this section. All of the set-theoretic claims in this section are elementary theorems of ZFCU. The details of the definitions and theorems (typically for the pure sets case) may be found in any good set theory text book; see for instance Jech (2003).

Functions are then defined as functional sets of pairs in the usual way. Given sets X and Y, a set f is said to be a function from X into Y, if (i) each element of f is a pair $\langle x, y\rangle$, with $x \in X$ and $y \in Y$; (ii) for each $x \in X$ there is a unique $y \in Y$ such that $\langle x, y\rangle$ is an element of f. More formally, let $f\colon X \to Y$ abbreviate:[39]

$$\forall z(z \in f \to \exists x \exists y(x \in X \wedge y \in Y \wedge z = \langle x, y\rangle)) \\ \wedge \forall x(x \in X \to \exists! y(y \in Y \wedge \langle x, y\rangle \in f))$$

Standard set theory may then be deployed to show that we have the functions needed to give a categorial semantics for $\mathcal{L}_{\beta,\in}$ under its intended interpretation over M. First we recap some elementary set theory.

Products and exponents. The Cartesian product of X and Y, $X \times Y$, and the set of all functions from X into Y, Y^X, are defined as follows

$$X \times Y =_{\text{df}} \{\langle x, y\rangle \mid x \in X, y \in Y\} \qquad Y^X =_{\text{df}} \{f \mid f\colon X \to Y\}$$

Whenever X and Y are sets, it is readily shown that $X \times Y$ and Y^X are also sets.

Granted that the universe M, the worlds W and the variables VAR each form a set, we may then construct for each category c a set M_c comprising all the category c intensions.

Intensions for basic and derived categories. First, note that the indices form a set $I = W \times M^{\text{VAR}}$. Pick distinct objects T and F to represent the two truth-values. The set of all category t-intensions, M_t, the set of all category e intensions, M_e, and for $d = c/(c_1, \ldots c_n)$, the set of all category d intensions, M_d, are defined recursively as follows: $M_t =_{\text{df}} \{T, F\}^I$, $M_e =_{\text{df}} M^I$, $M_d =_{\text{df}} M_c^{M_{c_1} \times \ldots \times M_{c_n}}$.[40]

With sets of indices and sets of each category of intension, it is then straightforward to specify which elements of the latter are the intensions of simple $\mathcal{L}_{\beta,\in}$ expressions. (See table 2.6.) Last of all, it remains to re-state the categorial compositionality principle in set-theoretic terms.

Compositionality. First define the set of *all* intensions based on the indices over the object universe, M_{sem}, and the set of all finite strings of intensions $\langle \boldsymbol{d}, \boldsymbol{c}_1, \ldots \boldsymbol{c}_n\rangle$

[39]The sentence in the text is not yet an $\mathcal{L}_{\beta,\in}$-sentence. We need further to eliminate the pair terms as follows: $z = \langle x, y\rangle$ may be taken as short for

$$\forall t\big(t \in z \leftrightarrow \forall u(u \in t \leftrightarrow u = x) \vee \forall u(u \in t \leftrightarrow u = x \vee u = y)\big)$$

and $\langle x, y\rangle \in f$ may be taken to abbreviate $\exists z(z = \langle x, y\rangle \wedge z \in f)$. As usual $\exists! y\phi$ abbreviates $\exists y(\phi \wedge \forall v(\phi[v/y] \to v = y))$ (picking v to avoid unintended variable binding).

[40]We rely here on the well-foundedness of the relative derivedness relation.

Table 2.6: Intensions encoded as set-functions

$$
\begin{aligned}
[\![v]\!] &= \{\langle i, \sigma_i(v)\rangle \in I \times M \mid i = \langle c, w, \sigma_i\rangle \text{ for some } c \in C,\ w \in W\} \\
[\![\beta]\!] &= \{\langle \boldsymbol{e}, \boldsymbol{t}\rangle \in M_e \times M_t \mid \boldsymbol{t}(i) = T \text{ iff } \beta(\boldsymbol{e}_1(i))\} \\
[\![\neg]\!] &= \{\langle \boldsymbol{t}_1, \boldsymbol{t}\rangle \in M_t \times M_t \mid \boldsymbol{t}(i) = T \text{ iff } \boldsymbol{t}_1(i) = F\} \\
[\![=]\!] &= \{\langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t}\rangle \in M_e \times M_e \times M_t \mid \boldsymbol{t}(i) = T \text{ iff } \boldsymbol{e}_1(i) = \boldsymbol{e}_2(i)\} \\
[\![\in]\!] &= \{\langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t}\rangle \in M_e \times M_e \times M_t \mid \boldsymbol{t}(i) = T \text{ iff } \boldsymbol{e}_1(i) \in \boldsymbol{e}_2(i)\} \\
[\![\wedge]\!] &= \{\langle \boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}\rangle \in M_t \times M_t \times M_t \mid \boldsymbol{t}(i) = T \text{ iff } \boldsymbol{t}_1(i) = T \text{ and } \boldsymbol{t}_2(i) = T\} \\
[\![\forall v]\!] &= \{\langle \boldsymbol{t}_1, \boldsymbol{t}\rangle \in M_t \times M_t \mid \boldsymbol{t}(i) = T \text{ iff } \boldsymbol{t}_1(i(v/a)) = T \text{ for every } a \in M\}
\end{aligned}
$$

(for $d = c/c_1, \ldots, c_n$), M_{seq}, as follows.

$$
\begin{aligned}
M_{\text{sem}} &= \bigcup\{M_c \mid c \in \text{CAT}\} \\
M_{\text{seq}} &= \bigcup\{M_d \times M_{c_1} \times \cdots \times M_{c_n} \mid d = c/c_1, \ldots c_n \in \text{CAT}\}
\end{aligned}
$$

Then, the composition-function, $\cdot\langle\!\langle \cdots \rangle\!\rangle$ which maps each string of intensions of the form $\langle \boldsymbol{d}, \boldsymbol{c}_1, \ldots \boldsymbol{c}_n\rangle$ (for $d = c/c_1, \ldots, c_n$) to the result of composing them, $\boldsymbol{d}(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)$, is defined as follows.

$$\cdot\langle\!\langle \cdots \rangle\!\rangle = \{\langle\langle \boldsymbol{d}, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\rangle, \boldsymbol{c}\rangle \in M_{\text{seq}} \times M_{\text{sem}} \mid \boldsymbol{c} = \boldsymbol{d}(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)\}$$

Finally the intension-function $[\![\cdot]\!]$, which maps each $\mathcal{L}_{\beta,\in}$-expression ε to its intension $[\![\varepsilon]\!]$, can be defined by recursion on complexity, via the clauses in table 2.6 and the following clause for complex expressions: $[\![\varepsilon(\varepsilon_1, \ldots, \varepsilon_n)]\!] = [\![\varepsilon]\!]\langle\!\langle [\![\varepsilon_1]\!], \ldots, [\![\varepsilon_n]\!] \rangle\!\rangle$.[41]

This completes the expansionist's encoding of categorial semantics for $\mathcal{L}_{\beta,\in}$ within standard set theory.

3.2 Absolutist semantics

The absolutist cannot always follow suit. Casting categorial semantics in standard first-order set theory is only an option when the members of the object language's universe form a set. In general, the absolutist claims, this may fail to be the case. In particular, if the quantifiers of the object language range over what he claims to be the ultimate, inexpandable universe $\boldsymbol{M}$ then, even when he deploys a metalanguage also interpreted over the ultimate universe, he cannot formulate a categorial semantics in the manner of the expansionist.

The trouble is this. When the members of the object language's universe $\boldsymbol{M}$ do

[41] We rely here on the well-foundedness of the relative-complexity relation.

not form a set, nor do the indices $\boldsymbol{I}$ based on this universe.[42] And without a set of indices, the bottom falls out of the set-theoretic construction of categorial semantics. According to the expansionist, category e intensions such as $[\![v]\!]$ are sets containing exactly one index-object pair for every index. But when the indices do not form a set, no such index-object pairs form a set.[43] Similar trouble arises for category t intensions. And without sets encoding category e and t intensions, the absolutist is also without n-tuples of these, let alone the *sets* of n-tuples of these, which the expansionist took to be the intensions of category t/e, $t/(e,e)$, t/t and $t/(t,t)$ expressions.

Still, what the expansionist achieves by drawing on an ontology surpassing that of the object language, the absolutist can achieve with extended ideological resources. In place of a hierarchy of sets built from members of the object language universe, a categorial semantics for $\mathcal{L}_{\text{ß},\in}$ may be framed in a metalanguage, interpreted over the same universe as the object language but deploying additional levels of ideology. Loose talk using English noun phrases like 'function' or 'intension' can be eliminated in favour of either higher-order or plural resources. We shall follow the latter option, although it is easily adapted to the former.

To implement it, we first enrich the metalanguage with suitable plural expressive resources. To start with let us add plural resources to the lexicon of $\mathcal{L}_{\text{ß},\in}$ to obtain the plural language of set theory with urelements, $\mathcal{L}^{pl}_{\text{ß},\in}$. This calls for a further basic category to join e and t: we add a category of plural terms: ee. The lexicon of the new language $\mathcal{L}^{pl}_{\text{ß},\in}$ then consists of that of $\mathcal{L}_{\text{ß},\in}$ together with the further expressions displayed in table 2.7. The intended interpretation of the new expressions may be informally glossed as follows. The plural quantifier $\forall vv$ may be read 'for any zero or more things...';[44] $x \prec \tau\tau$ may be read 'x is one of the objects $\tau\tau$'; $\tau\tau_1 \equiv \tau\tau_2$ may be read '$\tau\tau_1$ are the same things

[42] Proof: If the indices form a set $\boldsymbol{I}$, then member of the universe $\boldsymbol{M}$ form the set $\{\sigma(x_1) \mid \langle\sigma, w\rangle \in I\}$ (by Replacement) but *ex hypothesi* the members of the universe do not form a set

[43] Proof: If $\boldsymbol{e} = \{\langle i, a_i\rangle \mid i \in I\}$ is a set, then the indices form the set $\{i \mid \langle i, a_i\rangle \in \boldsymbol{e}\}$ (by Replacement). (More generally, the domain of a set-function is always a set.) But the indices do not form a set.

[44] This slight departure from the usual English meaning of 'any things' greatly simplifies the encoding, and will be attached to all plural quantifiers in the following. It is however inessential. See chapter 4 n. 8.

Table 2.7: The additional lexicon of $\mathcal{L}^{pl}_{ß,\in}$

Category ee.	plural variables: $xx, yy, zz, \ldots$
Category t/(e, ee).	the member-'plurality' predicate: $\prec$
Category t/(ee, ee)	the plural identity predicate: $\equiv$
Category ee/t.	plural term forming operators: λx, λy, λz, ...
Category t/t.	plural quantifiers: $\forall xx$, $\forall yy$, $\forall zz, \ldots$

as $\tau\tau_2$'; $\langle \lambda v.\phi \rangle$ may be read 'the zero or more things that satisfy ϕ'. (The λv-operator functions much like the abstraction operator $\hat{v}$ but delivers a plural term (category ee) rather than a predicate (category t/e).) What the expansionist is able to achieve with sets outside the universe, the absolutist can to some extent replicate with plural term denotations over the universe.[45]

Encoding functions as plural term denotations. Define Kuratowski pairs as before. Functions on members of the universe may be encoded as some functional pairs, severally. Given plural term denotations xx and yy, a plural term denotation *ff* is said to be a function from xx into yy, if (i) each member of *ff* is a pair $\langle x, y \rangle$, with $x \prec xx$ and $y \prec yy$; (ii) for each $x \prec xx$ there is a unique $y \prec yy$ such that $\langle x, y \rangle$ is a member of *ff*. More formally, let *ff*: $xx \to yy$ abbreviate:[46]

$$\forall z(z \prec ff \to \exists x \exists y(x \prec xx \land y \prec yy \land z = \langle x, y \rangle)) \\ \land \forall x(x \prec xx \to \exists! y(y \prec yy \land \langle x, y \rangle \prec ff))$$

This permits the absolutist to encode category t and e intensions. For example, $[\![v]\!]$ may be encoded as the plural term denotation comprising all index-object pairs $\langle i, \sigma_i(v) \rangle$. However, this approach will not permit us to encode the intensions of expressions from derived categories. Consider, for instance, the set predicate *ß*. Under the expansionist semantics this was encoded as a set whose elements were Kuratowski pairs of the form

[45]Plural or higher-order resources have often been called on in order to formulate absolutist-friendly formulations of model theory, whose models include one whose universe is comprehensive, and we shall borrow a number of the coding techniques from their presentations. See, for instance McGee (1997), Rayo and Uzquiano (1999), Rayo (2002), Williamson (2003), McGee (2006). The claims made about plural term denotations in the section are readily established in plural set theory, ZFCU^{pl}, in which the logical axioms are augmented with analogues of the usual axioms and rules for quantifiers and identity, and λ-exchange axioms of the form $u \prec \langle \lambda v.\phi \rangle \leftrightarrow \phi[u/v]$, and the Replacement schema is replaced with a single plural axiom (see, for instance, Uzquiano (2006, p. 309)).

[46]We apply the same abbreviations as in n. 39; moreover $\langle x, y \rangle \prec ff$ is short for $\exists z(z = \langle x, y \rangle \land z \prec ff)$. Such an encoding of functions is implicit in the definition of assignment and model employed in Rayo and Uzquiano (1999, p. 5) and explicit, for instance, in Uzquiano (2006, p. 309).

Table 2.8: The additional lexicon of $\mathcal{L}^{susupl}_{\text{ß},\in}$

Category eee.	Superplural variables: $xxx, yyy, zzz, \ldots$
Category t/(ee, eee).	the 'plurality'-'superplurality' predicate: $\prec_2$
Category t/(eee, eeee).	the 'superplurality'-'supersuperplurality' predicate: $\prec_3$
Category eee/t.	Superplural term forming operators: λxx, λyy, λzz, ...
Category eeee/t.	Supersuperplural term forming operators: λxxx, ...
Category t/t.	Superplural quantifiers: $\forall xxx$, $\forall yyy$, $\forall zzz, \ldots$

$\langle \boldsymbol{e}, \boldsymbol{t}\rangle$. But on the present approach, there are no such pairs: $\boldsymbol{e}$ and $\boldsymbol{t}$ are plural term denotations and pairs are always pairs of objects.

To work around this problem, the absolutist can call on yet further ideology, drawing on plural resources surpassing the plural. He may introduce superplural terms, that are to plural terms much as plural terms are to singular ones. Where a plural term denotes some objects, corresponding to a class of objects, a superplural term denotation corresponds to a 2-class, a class of classes of objects, or to some 1-class*es*. But the analogy cannot be taken too seriously. A term that denotes a class or a 2-class is a singular term, not a plural or superplural one; likewise a term that denotes some classes is a plural term not a superplural one.[47]

Anticipating a need for yet further ideology beyond superplural, we shall allow for two further categories of expression: *eee* for superplural terms, and *eeee* for supersuperplural terms, corresponding to 3-classes, classes of classes of classes of objects. We can then enrich the lexicon of $\mathcal{L}^{pl}_{\text{ß},\in}$ with the expressions in table 2.8 to obtain the supersuperplural language $\mathcal{L}^{susupl}_{\text{ß},\in}$. With these expressive resources, the absolutist has everything he needs to encode the intensions of all $\mathcal{L}_{\text{ß},\in}$ expressions.

[47]Rayo (2006) draws on superplural resources and beyond to develop a model theory for super^{n}plural languages, also extending the construction beyond the finite. See also Linnebo and Rayo (2012). The claims made about plural term denotations in this section are readily established in the plural set theory described in n. 45 enriched with analogues of the usual quantifier principles for superplural quantifiers, and λ-exchange principles for superplural and supersuperplural terms

$$uu \prec_2 \langle \lambda vv.\phi\rangle \leftrightarrow \phi[uu/vv]$$
$$uuu \prec_3 \langle \lambda vvv.\phi\rangle \leftrightarrow \phi[uuu/vvv]$$

Encoding functions as plural term denotations First we generalise pairing to plural term denotations as follows. Given plural term denotations xx and yy, we may define their '2-pair' $\langle xx, yy\rangle$ as follows.

$$\langle xx, yy\rangle =_{\text{df}} \langle \lambda z.\exists x(x \prec xx \wedge z = \langle 1, x\rangle) \vee \exists y(y \prec yy \wedge z = \langle 2, y\rangle)\rangle$$

This definition permits us to recover a plural version of the pairing theorem.[48]

$$\langle xx_1, yy_1\rangle = \langle xx_2, yy_2\rangle \text{ iff } xx_1 = xx_2 \text{ and } yy_1 = yy_2$$

2-pairs are plural term denotations. Consequently, functions mapping intensions, or other entities, encoded as plural term denotations, to others, may be encoded as superplural denotations. Much as before, we define fff: $xxx \to yyy$ to abbreviate:

$$\begin{aligned}&\forall zz(zz \prec_2 \mathit{fff} \to \exists xx \exists yy(xx \prec_2 xxx \wedge yy \prec_2 yyy \wedge zz = \langle xx, yy\rangle))\\ &\qquad\qquad \wedge \forall xx(xx \prec_2 xxx \to \exists! yy(yy \prec_2 yyy \wedge \langle xx, yy\rangle \prec_2 \mathit{fff}))\end{aligned}$$

Products and exponents revisited. Analogues of the earlier set theoretic definitions can now be given in plural and superplural terms. Define:

$$\begin{aligned} xx \times yy &=_{\text{df}} \langle \lambda z.(\exists x \prec xx)(\exists y \prec yy)(z = \langle x, y\rangle)\rangle\\ xxx \times yyy &=_{\text{df}} \langle \lambda zz.(\exists xx \prec_2 xxx)(\exists yy \prec_2 yyy)(zz = \langle xx, yy\rangle)\rangle\\ yy^{xx} &=_{\text{df}} \langle \lambda \mathit{ff}.\mathit{ff}\colon xx \to yy\rangle\\ yyy^{xxx} &=_{\text{df}} \langle \lambda \mathit{fff}.\mathit{fff}\colon xxx \to yyy\rangle \end{aligned}$$

Intensions for basic and derived categories. First encode the object language universe and the indices as plural term denotation $\boldsymbol{M} =_{\text{df}} \langle \lambda x.\exists y(y = x)\rangle$ and $\boldsymbol{I} =_{\text{df}} \langle \hat{x}.\exists w \exists \sigma(x = \langle w, \sigma\rangle)\rangle$. The definitions of $\boldsymbol{M}_c$ for $c = e$, t, t/t, $t/(t,t)$, t/e or $t/(e,e)$ then precede much as before, replacing set-theoretic products and exponents with their plural or superplural analogues as appropriate.

$$\begin{aligned} \boldsymbol{M}_t &=_{\text{df}} \langle \lambda x.(x = T \vee x = F)\rangle^{\boldsymbol{I}} & \boldsymbol{M}_e &=_{\text{df}} \boldsymbol{M}^{\boldsymbol{I}}\\ \boldsymbol{M}_{t/e} &=_{\text{df}} \boldsymbol{M}_t^{\boldsymbol{M}_e} & \boldsymbol{M}_{t/e,e} &=_{\text{df}} \boldsymbol{M}_t^{\boldsymbol{M}_e \times \boldsymbol{M}_e}\\ \boldsymbol{M}_{t/t} &=_{\text{df}} \boldsymbol{M}_t^{\boldsymbol{M}_t} & \boldsymbol{M}_{t/t,t} &=_{\text{df}} \boldsymbol{M}_t^{\boldsymbol{M}_t \times \boldsymbol{M}_t} \end{aligned}$$

Finally the intensions for the basic expressions of $\mathcal{L}_{ß,\in}$ are defined as in table 2.9.

Notice that unlike in the set-theoretic case, where we gave the analogous definitions for all categories of expressions, here we have only given the definitions for those categories of expression instantiated by simple $\mathcal{L}_{ß,\in}$ expressions. This marks an important difference between plural and set-theoretic resources for encoding functions. Whereas set-theoretic

[48] Linnebo and Rayo (2012) encode 2-pairs in essentially this way.

Table 2.9: Intensions encoded as higher-order term denotations

Write $\langle\lambda\langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t}\rangle.\phi(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t})\rangle$ to abbreviate the superplural term $\langle\lambda xx.\exists \boldsymbol{e}_1 \exists \boldsymbol{e}_2 \exists \boldsymbol{t}\big(\boldsymbol{e}_1 \prec \boldsymbol{M}_e \wedge \boldsymbol{e}_2 \prec \boldsymbol{M}_e \wedge \boldsymbol{t} \prec \boldsymbol{M}_t \wedge xx = \langle\langle \boldsymbol{e}_1, \boldsymbol{e}_2\rangle, \boldsymbol{t}\rangle \wedge \phi(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t})\big)\rangle$; similarly *mutatis mutandis* in the other cases.

$$
\begin{aligned}
[\![v]\!]^2 &=_{\mathrm{df}} \langle \lambda z.z = \langle i, \sigma_i(v)\rangle\rangle \\
[\![=]\!]^3 &=_{\mathrm{df}} \langle \lambda\langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t}\rangle.\forall i(\boldsymbol{t}(i) = T \leftrightarrow \boldsymbol{e}_1(i) = \boldsymbol{e}_2(i))\rangle \\
[\![\in]\!]^3 &=_{\mathrm{df}} \langle \lambda\langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{t}\rangle.\forall i(\boldsymbol{t}(i) = T \leftrightarrow \boldsymbol{e}_1(i) \in \boldsymbol{e}_2(i))\rangle \\
[\![\beta]\!]^3 &=_{\mathrm{df}} \langle \lambda\langle \boldsymbol{e}, \boldsymbol{t}\rangle.\forall i(\boldsymbol{t}(i) = T \leftrightarrow \beta(\boldsymbol{e}(i))\rangle \\
[\![\neg]\!]^3 &=_{\mathrm{df}} \langle \lambda\langle \boldsymbol{t}_1, \boldsymbol{t}\rangle.\forall i(\boldsymbol{t}(i) = T \leftrightarrow \boldsymbol{t}_1(i) = F)\rangle \\
[\![\wedge]\!]^3 &=_{\mathrm{df}} \langle \lambda\langle \boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}\rangle.\forall i(\boldsymbol{t}(i) = T \leftrightarrow (\boldsymbol{t}_1(i) = T \wedge \boldsymbol{t}_2(i) = T)\rangle \\
[\![\forall v]\!]^3 &=_{\mathrm{df}} \langle \lambda\langle \boldsymbol{t}_1, \boldsymbol{t}\rangle. \in M_t \times M_t \mid \forall i(\boldsymbol{t}(i) = T \leftrightarrow \forall a(a \prec \boldsymbol{M} \rightarrow \boldsymbol{t}_1(i(v/a)) = T))\rangle
\end{aligned}
$$

functions are type-free, functions encoded as plural terms denotations are typed. In the set-theoretic case all intensions are encoded as entities of the same type, namely objects. But in the plural case, as we come to intensions of more and more derived categories we require further and further plural resources. The intentions from categories t and e are encoded as plural term denotations, the intensions from categories t/e, $t/(e, e)$, t/t and $t/(t, t)$ as superplural term denotations. To encode intensions of categories higher up the derivedness ordering requires plural resources of greater and greater number, moving beyond supersuperplural, or super^{2}plural, to super^{3}plural, super^{4}plural, and so on.

This also poses a problem when it comes to stating compositionality. The expansionist was able to collect all intensions and all well-formed sequences of the same into single sets, and define the composition function as a subset of their Cartesian product. But, for the absolutist, the analogous procedure founders at the first step. There is no plural or superplural term denotation to be the analogue of the set of all intensions. Even if we were to add super^{n}plural terms for each finite n, no super^{n}plural term denotation will encompass all intensions, for some of these will need to be encoded as super^{m}plural term denotations, where m exceeds n.

One response to this problem is to tone down the compositionality principle.[49]

[49] A second approach would be to tone up the ideological resources. The absolutist could follow Linnebo and Rayo (2012) in employing plural variables of transfinite number, allowing what we might call super$^\omega$plural term denotations to encompass super^{n}plural term denotations of any finite number. But I shall leave the absolutist to work out the details of this approach.

Rather than allowing for the composition of arbitrary intensions, the absolutist can cap the compositionality principle at some fixed level, suitable for the language at hand. For $\mathcal{L}_{ß,\in}$ this is easily achieved. We want a function $\langle\!\langle \cdot \rangle\!\rangle$ that maps the intensions of the immediate constituents of each complex expression in $\mathcal{L}_{ß,\in}$ to the intension of the complex expression. In $\mathcal{L}_{ß,\in}$ complex expressions are always formed by composing a predicate, connective, or quantifier—whose intensions are all encoded by superplural term denotations—with one or two variables, or sentences—whose intensions, and pairs of the same, are all encoded by plural term denotations. Consequently the concatenation function only need deal with sequences of the form $\langle \boldsymbol{d}, \boldsymbol{c} \rangle$ and $\langle \boldsymbol{d}, \langle \boldsymbol{c}_1, \boldsymbol{c}_2 \rangle \rangle$ as input, where $\boldsymbol{c}$, $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ are intensions with category t or e and $\boldsymbol{d}$ is an intensions with category t/e, $t/(e,e)$, t/t and $t/(t,t)$. Employing a similar coding trick to before to encode mixed '3-pairs', such as $\langle xxx, yy \rangle$ as superplural term denotations such a function may be encoded as a supersuperplural term denotation.

Compositionality. First define mixed 2- and 3-pairs as follows.

$$\begin{aligned} \langle x, yy \rangle &=_{\mathrm{df}} \langle \lambda z. \exists y \prec yy (z = \langle 1, x \rangle \vee z = \langle 2, y \rangle) \rangle \\ \langle xxx, yy \rangle &=_{\mathrm{df}} \langle \lambda zz. \exists xx \prec xxx (zz = \langle \langle \lambda t.t = 1 \rangle, xx \rangle \vee zz = \langle \langle \lambda t.t = 2 \rangle, yy \rangle) \rangle \end{aligned}$$

These are readily seen to conform to the pairing principle. Functions mapping objects to plural term denotations may be defined in the obvious: $fff : xx \to yyy$ abbreviates:

$$\forall zz(z \prec_2 fff \to \exists x \exists yy(x \prec_2 xx \wedge yy \prec_2 yyy \wedge zz = \langle x, yy \rangle)) \\ \wedge \forall x(x \prec xx \to \exists! yy(yy \prec_2 yyy \wedge \langle x, yy \rangle \prec_2 fff))$$

The composition function may then be defined as follows. Let p^n abbreviate t/e^n (as above) and q^n abbreviate t/t^n; and write $\exists \boldsymbol{c}\phi$ to abbreviate the result of replacing $\boldsymbol{c}$ for a suitable variable in $\exists \boldsymbol{c}(\boldsymbol{c} \prec \boldsymbol{M}_c \wedge \phi))$. $\langle\!\langle \cdot \rangle\!\rangle$ is encoded by the following supersuperplural term denotation.

$$\begin{aligned} \langle \lambda zzz. \big(& \exists xxx \exists xx \exists yy \, zzz = \langle \langle xxx, xx \rangle, yy \rangle \\ & \wedge \big(\exists \boldsymbol{p}^1 \exists \boldsymbol{e}(xxx = \boldsymbol{p}^1 \wedge xx = \boldsymbol{e} \wedge yy = \boldsymbol{p}^1(\boldsymbol{e})) \\ & \quad \vee \exists \boldsymbol{p}^2 \exists \boldsymbol{e}_1 \exists \boldsymbol{e}_2 (xxx = \boldsymbol{p}^2 \wedge xx = \langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle \wedge yy = \boldsymbol{p}^2(\boldsymbol{e}_1, \boldsymbol{e}_2)) \\ & \quad \vee \exists \boldsymbol{q}^1 \exists \boldsymbol{t}(xxx = \boldsymbol{q}^1 \wedge xx = \boldsymbol{t} \wedge yy = \boldsymbol{q}^1(\boldsymbol{t})) \\ & \quad \vee \exists \boldsymbol{q}^2 \exists \boldsymbol{t}_1 \exists \boldsymbol{t}_2 (xxx = \boldsymbol{q}^2 \wedge xx = \langle \boldsymbol{t}_1, \boldsymbol{t}_2 \rangle \wedge yy = \boldsymbol{q}^2(\boldsymbol{t}_1, \boldsymbol{t}_2))\big)\big) \rangle \end{aligned}$$

Finally, since all complex expressions in $\mathcal{L}_{ß,\in}$ are sentences, the intension function $[\![\cdot]\!]$ may be encoded as a superplural term denotation mapping sentences (objects) to category t intensions (plural term denotations). We employ Frege's trick to define $[\![\cdot]\!]$ as the smallest such superplural term denotation that satisfies the following conditions.

$$F(\text{ß}(v)) = [\![\text{ß}]\!]\langle\!\langle[\![v]\!]\rangle\!\rangle \qquad F(u \in v) = [\![\in]\!]\langle\!\langle[\![u]\!],[\![v]\!]\rangle\!\rangle \qquad F(u = v) = [\![=]\!]\langle\!\langle[\![u]\!],[\![v]\!]\rangle\!\rangle$$
$$F(\neg\phi) = [\![\neg]\!]\langle\!\langle F(\phi)\rangle\!\rangle \quad F(\phi \wedge \psi) = [\![\wedge]\!]\langle\!\langle F(\phi), F(\psi)\rangle\!\rangle \quad F(\forall v\phi) = [\![\forall v]\!]\langle\!\langle F(\phi)\rangle\!\rangle$$

Let $\textsc{int}(F)$ be the result of prefixing the conjunction of these conditions with the quantifiers $(\forall\phi \in \textsc{sent})(\forall\psi \in \textsc{sent})$. Then we define $[\![\cdot]\!]$ as follows:[50]

$$\langle\lambda xx.(\exists\phi \in \textsc{sent})\exists \boldsymbol{t}(xx = \langle\phi, \boldsymbol{t}\rangle \wedge \forall F(F\colon \textsc{sent} \to M_t \wedge \textsc{int}(F) \to \langle\phi, \boldsymbol{t}\rangle \prec_2 F))$$

This completes the plural encoding of categorial semantics on behalf of the absolutist. Of course, the absolutist will need to take the informal part of the exposition here with a large pinch of salt. In attempting to motivate the statement of the theory in $\mathcal{L}^{susupl}_{\text{ß},\in}$, I have used English quantifier expressions like 'every superplural term denotation' or 'some category t intension' and these do no better than 'every concept' or 'some plurality' in quantifying over anything other than *objects*. If there are such *things* as superplural term denotations or pluralities or category t intensions then they are objects.[51]

Given a suitably liberal metalanguage, the expansionist can take quantifiers like 'every category t intension' at face value, ranging over objects—indeed, she might claim, sets of index-object pairs—and view the exposition as an account of how such ontology *outside* the object language universe can be replaced with additional ideology *over* it. But for the absolutist, in the crucial case when he claims the universe to be ultimate, no such perspective is available. The informal exposition in English must be viewed as a means to *communicate* what can be *said* only with sufficient plural (or higher-order) resources.

3.3 Restrictionist semantics

Neither the expansionist's nor the absolutist's encoding strategies can be followed by the restrictionist to formulate categorial semantics. The root of the problem is at the very centre of (anti-expansionist) restrictionism: the universe may be ultimate, but no

[50] Compare Rayo and Uzquiano (1999, p. 9).
[51] Compare Williamson (2003, p. 420).

domain encompasses the entire universe. Consequently, the categorial semantic theory for a language with the ultimate universe $\boldsymbol{M}$—and the attendant theory of functions—must be framed in a meta-context $C^\star$, or more generally a meta-perspective, whose quantifiers range only over $D_{C^\star}$.

This places severe limitations on what functions are available, regardless of what resources we might deploy to encode them. Working within $D_{C^\star}$, any set-function deployed must be a member of this domain; and so cannot have the domain, let alone the ultimate universe $\boldsymbol{M}$, as its domain or range. Following the absolutist, any function *over* $D_{C^\star}$ can be encoded using a plural term denotation applying to pairs. But, since the pairs available in $C^\star$ are pairs of members of $D_{C^\star}$, the domain and range of any function encoded by such pairs is contained within $D_{C^\star}$. Helping ourselves to further ideological resources enables us to encode further functions of functions over $D_{C^\star}$, as above, but does not enable us to break free from $D_{C^\star}$ at the lowest level. A 2-pair of plural term denotations $\langle xx, yy\rangle$ is available from the perspective of $D_{C^\star}$ only if its co-ordinates, xx and yy, only denote members of $D_{C^\star}$. Consequently, a category t/t intension encoded by a third-order term denotation is available in $C^\star$ only if each category t intension in its domain is also available from this perspective. Similarly, for the other categories.

Without arbitrary functions over the object language universe $\boldsymbol{M}$, the intensions deployed by the categorial semantics outlined above are simply missing. Consider the intension of $\forall v$. As in section 2, two options are apparent according to whether the restriction maintains that contextual restriction is solely due to local restrictions, or incorporates indexicality in determiners via shifts in the background domain.

$$\big(\llbracket\forall v\rrbracket_C(\boldsymbol{t})\big)(i) = T \text{ iff } \boldsymbol{t}(i[v/a]) = T \text{ for every } a \in \boldsymbol{M}$$
$$\big(\llbracket\forall v\rrbracket_C^{D_C}(\boldsymbol{t})\big)(i) = T \text{ iff } \boldsymbol{t}(i[v/a]) = T \text{ for every } a \in D_C$$

Setting aside the philosophical difficulties facing the first approach, the intension deployed by the absolutist is in any case unavailable to the restrictionist relative to any perspective. This function is available in $C^\star$ only if each category t intension $\boldsymbol{t}$ in its domain is available in $C^\star$; $\boldsymbol{t}$ is available only if each index $\langle\sigma, w\rangle$ over $\boldsymbol{M}$ is available;

$\langle \sigma, w\rangle$ is available only if σ is available; and σ is available only if each object $\sigma(v)$ in its range is a member of $D_{C^\star}$. But in general this fails to be the case: σ is an arbitrary assignment over the universe $\boldsymbol{M}$, not the limited theoretical domain $D_{C^\star}$. The second candidate intension is likewise unavailable to the restrictionist for the very same reason: the restriction to M_c in the defining condition affects the function's values but has no effect on the extent of its domain, which remains too extensive for $C^\star$.

Instead, the best the restrictionist can do is to restrict the intension of the quantifier to category t intensions based on indices over the limited, metatheoretic domain, replacing $\boldsymbol{M}$ with $D_{C^\star}$ from the very outset. The indices I must be replaced with the indices $I^\star$ that are available in the context, namely those indices $i^\star$ whose assignment $\sigma_i^\star$ maps each variable to a member of $D_{C^\star}$. Similarly $\boldsymbol{M}_t$ must be replaced with $M_t^\star$ comprising the category t intensions that are available in the context, namely functions $\boldsymbol{t}^\star$ from $I^\star$ to $\{T, F\}$, and so on.

The putative intensions for the quantifier based on the universe $\boldsymbol{M}$ must be replaced by analogues based on the theoretical domain $D_{C^\star}$.

$$\big(\llbracket \forall v \rrbracket_C^{D_{C^\star}}(\boldsymbol{t}^\star)\big)(i^\star) = T \text{ iff } \boldsymbol{t}^\star(i^\star[v/a]) = T \text{ for every } a \in D_C$$

$$\big(\llbracket \forall v \rrbracket_C^{D_C \cap D_{C^\star}}(\boldsymbol{t}^\star)\big)(i^\star) = T \text{ iff } \boldsymbol{t}^\star(i^\star[v/a]) = T \text{ for every } a \in D_C \cap D_{C^\star}$$

The trouble with this approach is that it fails to yield the correct truth-conditions for sentences of the form $\forall v\phi$ in contexts wider than the theoretical context $C^\star$. Indeed, since in this case, the intension of ϕ is unavailable in $C^\star$, it delivers no truth-conditions at all. Similar problems arise in the extensional setting. The extension of a quantifier in $C^\star$ can only comprise predicate extensions that are available from the perspective of that context.

4 Philosophical Assessment

The problem faced by the restrictionist when it comes to stating categorial and extensional semantic theories is at bottom the same problem she faces in framing Tarskian truth-conditions. We can frame an adequate semantic theory for a language relative to a given context C, delivering the correct truth-conditions for each sentence as then interpreted, only if the theoretical perspective $C^\star$ is at least as wide as C. But when the object language is interpreted over the ultimate universe—as the restrictionist allows it may be—there is no meta-perspective that exceeds every object perspective. So far then the charge of semantic pessimism would appear to stand firm.

Semantic pessimism for contextual restrictionists. The contextual restrictionist *cannot* (by her lights) give adequate semantic theories for languages that deploy restricted quantifiers in a suitable metalanguage.

At best, the contextual restrictionist can give partial semantic theories determining the behaviour of the language in a restricted range of contexts. Picking a sufficiently liberal meta-context, for instance, the restrictionist can adapt expansionist semantics for $\mathcal{L}_{ß,\in}$, to yield the theory of the interpretation of this language $\mathcal{L}_{ß,\in}$ at some fixed context. But this move is a means of learning to live with semantic pessimism rather than avoiding it. The units of semantic theories are languages, not contexts. The first desideratum of semantic theorising remains unmet: the restrictionist is unable to give a semantic theory that yields the correct truth-conditions for each sentence of $\mathcal{L}_{ß,\in}$ *in each context.* The same is true of other languages interpreted over the ultimate universe.

For the absolutist, however, the shift from Tarskian to extensional and categorial frameworks makes a difference. In the Tarskian setting, the absolutist had no trouble stating semantic clauses for purportedly absolutely general quantifiers; in the categorial setting he seems to fare rather less well than his expansionist opponent.

The absolutist-friendly supersuperplural semantic theory for the language of set theory $\mathcal{L}_{ß,\in}$ outlined above scores worse than its first-order expansionist analogue in two respects. In the first place, without a move to infinitely superplural expressions, the absolutist cannot formulate the principle of compositionality in full generality but must instead make do with a curtailed version, only dealing with finitely many categories of intension. In the second place, even with compositionality cut back only to apply to the categories of intension instantiated by $\mathcal{L}_{ß,\in}$-expressions, the absolutist semantics utilises substantially greater ideological resources than its expansionist analogue. The categorial semantics for the *first*-order language deploys *superplural* quantification and *supersuperplural* terms.

Can the absolutist avoid these worries with a less ideologically profligate encoding of intensions? Intensions have been encoded as plural and superplural term denotations. We have seen that they cannot be set-functions, but the absolutist might still look for yet other *objects* of $\boldsymbol{M}$, perhaps class- or 2-class-functions, or similar, to do their work.

Consequently, the absolutist might seek to interpret the higher-order theory—or at least enough of it for the purpose at hand—in a first-order theory. To start with, for category t intensions, he could seek a mapping $\boldsymbol{t} \mapsto \hat{\boldsymbol{t}}$, which maps each category t intension, encoded as a plural term denotation $\boldsymbol{t}$, to an *object* $\hat{\boldsymbol{t}}$ in $\boldsymbol{M}$ which encodes it; together with a predicate $\phi(i, \mathrm{V}, \hat{\boldsymbol{t}})$ encoding that the intension $\boldsymbol{t}$ maps the index i to the truth-value V

The expansionist *can* do this. Each plural term denotation comprising pairs of members of $\boldsymbol{M}$ is mapped to the corresponding set with these pairs as element (some of which fall outside $\boldsymbol{M}$) and $\phi(i, \mathrm{V}, \hat{\boldsymbol{t}})$ is defined as $\langle i, \mathrm{V} \rangle \in \hat{\boldsymbol{t}}$. But no such approach can succeed if each object $\hat{\boldsymbol{t}}$ encoding the intension $\boldsymbol{t}$ is confined to falling within the object language universe $\boldsymbol{M}$, as must be the case if $\boldsymbol{M}$ is the ultimate universe. The reason is simple. There are more category t intensions than members of $\boldsymbol{M}$. For there are as many indices as objects and, by an application of Cantor's diagonal argument, there are more functions from indices to truth-values than there are indices. It follows that there are too many intensions to be uniformly encoded in this way (since $\boldsymbol{t} \mapsto \hat{\boldsymbol{t}}$ would be an injective mapping from category t intensions to objects). Consequently, to generalise about *every* category t-intension, singular first-order quantification over $\boldsymbol{M}$ will not suffice. Without further ontology available, the absolutist must call on further ideology.

This will not come as too much of a surprise to seasoned absolutists. For it is well known that quantification over a given universe is ill-suited for theorising about *every* possible interpretation of a lexicon over that universe. This much is one of the lessons of Williamson's take on the Russell paradox. Williamson focuses on the (extensional) interpretation of a single predicate letter P. Assuming only that interpretations are themselves members of the purportedly inexpandable universe $\boldsymbol{M}$, there is no hope for the thesis that, for any predicate ϕ, there is an interpretation y of P under which the predicate letter P applies to every ϕ and nothing else. For writing 'P applies to x under y' as $A_P(x, y)$, this thesis amount to the following schema:

$$\exists y \forall x (A_P(x, y) \leftrightarrow \phi).$$

And treating $\exists y$ as a first-order quantifier, this is nothing more than a notational variant on the Naïve Comprehension Schema for sets, generating a contradiction when we substitute $\neg A_p(x, x)$ for ϕ. In order then to allow any predicate ϕ to determine a possible interpretation of P, we cannot treat $\exists y$ as quantifying over the objects in the domain of $\forall x$. Since Williamson takes this to encompass absolutely everything, he proposes a second-order construal of such interpretations (See Williamson (2003, §IV).)

Accordingly, certain sorts of theorising about first-order languages must be conducted in a second-order or plural language. In particular, we must appeal to such ideology in order to engage in what Linnebo and Rayo (2012) dub 'generalized semantics', where we are interested in *arbitrary* reinterpretation of the language's expressions. For example, following Rayo, let us call a model theory for a language 'strictly adequate' if every possible extension of an expression is assigned to that expression in some model. By the absolutist's lights, we cannot give a strictly adequate model theory for the language of set theory $\mathcal{L}_{\beta,\in}$ in set theory itself. If we identify extensions with sets we ignore those extensions—including the intended extensions of $\in$, $=$ and β over the ultimate universe—that are not set-sized. And Rayo observes that since any plural term denotation could encode a possible extension, and there are more plural term denotations than objects, no model theory framed in a first-order language quantifying only over objects can be strictly adequate. We must instead draw on plural resources. (See Rayo (2006, §9.7))

But once we take categorial semantic theories into account, a further point is apparent. It is not just 'generalized semantics' that requires non-first-order resources. The ability to quantify over arbitrary intensions is needed to give a categorial semantic theory for a *single* interpretation of a *particular* lexicon. In particular, it is needed to give a categorial semantics for the language of set theory $\mathcal{L}_{\text{ß},\in}$ under what by the absolutist's lights is its unique intended interpretation over the ultimate universe. In order to state the compositionality principle—even when restricted to the semantic categories instantiated by $\mathcal{L}_{\text{ß},\in}$ expressions—we need to generalise *inter alia* about arbitrary category t intensions over $\boldsymbol{M}$, and this requires plural resources.

The need for further plural resources extends in two dimensions. First, it generalises across to the extensional semantics outlined for $\mathcal{L}_{\text{GQ}}$. Arbitrary category t extensions—truth-values—are unproblematic, but there are too many unary predicate extensions over $\boldsymbol{M}$, which may simply be taken to be plural term denotations, for these to be uniformly encoded as members of M. Consequently, the absolutist must once again deploy second-order or plural quantification in order to formulate the compositionality clauses in the extensional semantics.

Second, in each case, further and further resources are required as we ascend the hierarchy of categories. In the extensional case, since there are as many predicate extensions (category t/e) as there are plural term denotations, and there are more quantifier extensions (category $t/(t/e)$) than predicate extensions, we can do no better than superplural term denotations to uniformly encode quantifier extensions.[52] In the intensional case, the fact that intensions are functions from indices to extensions calls for one further degree of pluralness. There are as many category t intensions as there are plural term denotations. And there are more category t/e intensions—functions mapping category e-intensions to category t intensions—than there are category t intensions. Consequently to uniformly encode arbitrary category t/e intensions we can do no better than superplural term denotations, as employed above. Similarly, there are as many category t/e intensions as there are superplural term denotations. And there are more category $t/(t/e)$ intensions than category t/e intensions. Consequently to uniformly encode arbitrary category $t/(t/e)$ intensions we can do no better than supersuperplural term denotations. And so on.

[52] Compare 'Semantic Ascent' in Rayo (2006, p. 244).

Extensional and categorial approaches to semantics require the absolutist to take on a great deal more ideology than his expansionist opponent. In particular, this leads to a *restricted* form of semantic pessimism for the prospects of carrying out certain forms of semantic theorising in a *first-order* metalanguage.

Categorial and extensional pessimism for absolutists deploying first-order metalanguages.

> The absolutist *cannot* (by his lights) give adequate categorial or extensional semantic theories for languages that deploy *absolutely* general quantifiers in a suitable (non-plural) first-order metalanguage.

The absolutist may object that this conclusion has been reached on the basis of premisses that have been whistled rather than stated. No sense—he may claim—can be made of sentences like 'There are more category t intensions than members of $\boldsymbol{M}$' apart from as trivially false claims. For, as in the introduction, he has been instructed to read singular talk of universes, in plural terms. When he takes $\boldsymbol{M}$ to be the ultimate domain, the sentence is thus elliptical for 'There are more category t intensions than there are objects'. And this is trivially false. If there are any such things as category t intensions, then they are objects; and some of the objects cannot be more numerous than *all* of the objects.

As before, the expansionist *can* make good sense of such claims at face value from the perspective of a sufficiently liberal metalanguage. In this case, the cardinality claim can be established within first-order set theory, by showing that there is no injective function with $\boldsymbol{M}$ as its domain and only t-intensions—construed as set-functions—in its range. But the absolutist may justly complain that if *this* is the intended reading, then the argument for pessimism relies on a premiss that will only be true if we presuppose that his view fails. Moreover, when—as he maintains—the members of the object language universe do not form a set, such set-theoretic arguments tell us nothing about how many objects there are compared to intensions: there is no injective function with $\boldsymbol{M}$ as its domain *full stop*; since the domain of a set-function is always a set.

In response, to avoid begging the question against absolutism, the relativist must suspend the assumption of a more liberal metalanguage, and make the argument in plural terms. The conclusion, the analogue of the English sentence: 'Category t intensions cannot be uniformly encoded as objects', is treated as schematic:

No uniform encoding. $\neg\exists \mathit{fff}\big(\mathit{fff}\colon \boldsymbol{M}_t \to \boldsymbol{M} \wedge \forall \boldsymbol{t}\forall i\big(\boldsymbol{t}(i) = T \leftrightarrow \phi(i, T, \mathit{fff}(\boldsymbol{t}))\big)\big)$

This claim may be proved in superplural set theory (see notes 45 and 47). The conclusion is reached in two steps. First we prove a superplural version of Cantor's theorem, which regiments the English sentence 'There are more category t intensions than indices' in a manner analogous to the standard set-theoretic rendering of this comparative cardinality claim as 'There is no injective function mapping category t intensions to indices'.

Cantor's theorem $\neg\exists ggg(ggg\colon \boldsymbol{M}_t \to \boldsymbol{I} \wedge \forall i\forall i'(ggg(i) = ggg(i') \to i = i')$

This may be proved with the usual sort of diagonal construction.[53]

Proof sketch. We reason informally in superplural set theory. Suppose for *reductio* that there is $ggg : ggg\colon \boldsymbol{M}_t \to \boldsymbol{I}$, injective as in the statement. For $\boldsymbol{t} \prec_2 \boldsymbol{M}_t$ write $\bar{\boldsymbol{t}}$ for the unique index i such that $\langle \boldsymbol{t}, i\rangle \in ggg$. Consider

$$rr = \langle \lambda x.(\exists i \prec \boldsymbol{I})\big((x = \langle i, T\rangle \wedge \exists \boldsymbol{t}(\bar{\boldsymbol{t}} = i \wedge \boldsymbol{t}(\hat{\boldsymbol{t}}) = F) \vee (x = \langle i, F\rangle \wedge \neg\exists \boldsymbol{t}(\bar{\boldsymbol{t}} = i \wedge \boldsymbol{t}(\bar{\boldsymbol{t}}) = F))\big)\rangle$$

Clearly $rr \prec_2 \boldsymbol{M}_t$. Consider $rr(\overline{rr})$. Suppose $rr(\overline{rr}) = T$. Then, for some $\boldsymbol{t}$, $\bar{\boldsymbol{t}} = \overline{rr}$ and $\boldsymbol{t}(\bar{\boldsymbol{t}}) = F$; so, by the injectiveness of ggg, $\boldsymbol{t} = rr$ and $rr(\overline{rr}) = F$. But now given that $rr(\overline{rr}) = F$, there is some $\boldsymbol{t}$ (namely, rr) such that $\bar{\boldsymbol{t}} = \overline{rr}$ and $rr(\overline{rr}) = F$; consequently, $rr(\overline{rr}) = T$. Contradiction. □

Second since there are as many indices as objects, we can exhibit an injective function from object to indices

Proof sketch of No Uniform Encoding. Suppose for reductio that fff provides the uniform encoding as in the statement with respect to ϕ. For $\boldsymbol{t} \prec_2 \boldsymbol{M}_t$ write $\hat{\boldsymbol{t}}$ for the unique object such that $\langle \boldsymbol{t}, i\rangle \in \mathit{fff}$. Notice that fff is injective: suppose $\hat{\boldsymbol{t}_1} = \hat{\boldsymbol{t}_2}$; then for any index i, $\boldsymbol{t}_1(i) = T$ iff $\phi(i, T, \hat{\boldsymbol{t}_1})$ iff $\phi(i, T, \hat{\boldsymbol{t}_2})$ iff $\boldsymbol{t}_1(i) = T$; so $\boldsymbol{t}_1 = \boldsymbol{t}_2$. Second, consider

$$ggg =_{\mathrm{df}} \langle \lambda zz.\exists \boldsymbol{t}(zz = \langle \boldsymbol{t}, \langle w_o, \sigma_0[\hat{\boldsymbol{t}}/x_1]\rangle\rangle$$

Clearly $ggg\colon \boldsymbol{M}_t \to I$. Moreover, ggg is injective. If $\langle w_o, \sigma_o[\hat{\boldsymbol{t}_1}/x_1]\rangle = \langle w_o, \sigma_o[\hat{\boldsymbol{t}_2}/x_1]\rangle$; then $\hat{\boldsymbol{t}_1} = \sigma_o[\hat{\boldsymbol{t}_1}/x_1](x_1) = \sigma_o[\hat{\boldsymbol{t}_2}/x_1](x_1) = \hat{\boldsymbol{t}_2}$; so, by the injectiveness of fff, $\boldsymbol{t}_1 = \boldsymbol{t}_2$. This contradicts Cantor's theorem. □

[53]Shapiro offers a rather different encoding of a second-version of Cantor's theorem. (1991, thm. 5.3). Linnebo and Rayo (2012) sketch a proof of a version more like the present one.

The cardinality objection can be made without presupposing relativism. The difficulties in elucidating superplural theories and the extra logical resources needed to demonstrate theses in such theories, however, also highlight a further cost for the absolutist. For if, as it seems to, English lacks the superplural or higher-order quantification required to frame a categorial semantics for $\mathcal{L}_{\beta,\in}$ or an extensional semantics for $\mathcal{L}_{\mathrm{GQ}}$, absolutism engenders a more troubling version of semantic pessimism.

Categorial and extensional pessimism for absolutists theorising in English. The absolutist *cannot* (by his lights) give adequate categorial or extensional semantic theories for languages that deploy *absolutely* general quantifiers in English.

This conclusion is more worrying since it commits the absolutist to claiming that a substantial strand of empirical semantics is deploying a defective methodology. Barwise and Cooper are just two examples of linguists who attempt to uncover the semantic properties of English determiners, working with extensional semantics against a background of first-order set theory.[54] If English is an absolutely general language, as absolutists have tended to claim, and the superplural resources required generalise about arbitrary quantifier extensions are unavailable in English, this approach misfires. Consider for example the hypothesis that all natural language quantifiers of the form $\det(\theta_1)$ live on the extension of θ_1. Linguists regard this hypothesis as well-confirmed, yet if it is stated in a version of English—as it invariably is—absolutists must reject it as trivially false. On this view, in unrestricted contexts, quantifiers like EVERY(THING) live on no set-sized domain.

The pessimistic conclusion is tentative since it is controversial whether or not English does have the requisite expressive resources.[55] But the availability of such resources does little to remove the commitment to pessimism about semantic theorising as it is actually carried out. Even if linguists do speak a language which has the resources to correctly encode semantic theories for absolutely general languages, they persist in misusing it, deploying singular terms like 'predicate extension' and 'quantifier denotation'

[54] See also, for instance, Keenan (1984), Keenan and Stavi (1986), Westerståhl (1985b), and Peters and Westerståhl (2006).

[55] Linnebo and Nicolas (2008) argue for the availability of superplural quantification in English.

when they should be using plural and superplural ones. The theories they actually give contain many trivial falsehoods.

This is far from a decisive objection against absolutism. In the first place, absolutists always have the Tarskian semantics to fall back on. In the second, the widespread error amongst linguists may be open to a reasonable explanation, and the absolutist may claim that the distinctively linguistic insights can be easily transposed to a plural or higher-order setting, purging them of the misguided assumption of a set-sized universe. Still, I submit that it goes to show that in the categorial and extensional setting absolutism leads to an uncomfortable if not insurmountable amount of semantic pessimism.

Expansionists, on the other hand, can accept the linguist's assumption of a set universe as perfectly legitimate in all cases. No matter what the object language is there is always another version of English suitable to function as the metalanguage. Unlike restrictionism, this view is quite compatible with a robust version of semantic optimism.

Semantic optimism for expansionists. The expansionist *can* (by her lights) give adequate categorial or extensional semantic theories for languages that deploy unrestricted quantifiers over a non-ultimate universe, including versions of English, in other versions of English.

Worries about semantic pessimism, then, do not stand up against relativism in general. Once we have distinguished between restrictionist and expansionist variants of relativism, we can see that not only does expansionism avoid the pessimism faced by the restrictionist, but it also fares rather better than absolutism when it comes to stating certain styles of semantic theory in the standard manner in English.

3 Expressive Issues

1 The Statement Problem

David Lewis offers a short response to those who attempt to adopt relativism about quantifiers in response to philosophical difficulties:

> Maybe [he] replies that some mystical censor stops us from quantifying over absolutely everything without restriction. Lo, he violates his own stricture in the very act of proclaiming it. (Lewis 1991, p. 68.)

This one line repudiation of relativism is not the only example of an eminent philosopher curtly rejecting this view on the grounds that the relativist is unable to successfully state it. Vann McGee sets aside relativism as 'not a serious worry' in two short paragraphs.

> The reason [relativism] is not a serious worry is that the thesis that, for any discussion, there are things that lie outside the universe of discourse of that discussion is a position that cannot be coherently maintained. Consider the discussion we are having right now. We cannot coherently claim that there are things that lie outside the universe of our discussion, for any witness to the truth of that claim would have to lie outside the claim's universe of discourse.
>
> Of course, the fact that a thesis cannot be coherently maintained does not strictly entail that the thesis is false. Even so, the fact that we cannot coherently hold the theory is surely reason enough not to embrace it. (McGee 2000, p. 55.)

The problem has also been pressed and examined in considerably more detail by Timothy Williamson, who concludes that '[g]enerality-relativists seem to be unable to articulate their position in any coherent way' (2003, p. 433).

1.1 Failed attempts

To elaborate on the problem let us consider three attempts to state relativism. Suppose that the contextual restrictionist attempts to state her view by uttering any of the following:

(1) Not everything is quantified over in any context.

(2) For any context C, there is a context C' that countenances things which C does not, but not vice versa.

(3) No domain has everything as a member.
$\forall xx \neg \forall x (x \prec xx)$

The second attempt here is a somewhat simplified English paraphrase of a semi-formal statement presented in print by Geoffrey Hellman (2006) in response to Williamson's objection.[1]

Each of these three attempts must be made in a context (and in a language): label the contexts respectively C_1, C_2 and C_3 (leaving the language implicit).[2] Now if the first attempt is to succeed the range of the quantifier 'any context' must encompass every context, in particular, C_1. Likewise the range of the quantifier 'any context' in C_2 must encompass C_2; and the range of the quantifier 'no domain' in C_3 must encompass the domain—D_3, call it—of C_3.[3] Consequently, given that the context or domain is in the range of the relevant quantifier, the first attempted general statement of relativism is

[1]His actual formulation is this

$$\forall C \exists C' (\text{`the range, } R'\text{, of q's } \textit{countenanced} \text{ in } C' \text{ exceeds that, } R\text{, } \textit{countenanced} \text{ in } C \\ \& \text{ commitments in } C' \text{ are well motivated relative to } C\text{'})$$

where 'q's' is short for quantifiers and Hellman explains: 'Here 'exceeds' means that C' countenances things which C does not, but not vice versa, however this is spelled out precisely' (Hellman 2006, p. 77) Nothing, so far as I can see, is lost by the omission of reference to ranges in the simpler formulation given in the text here, and this seems to fit better with Hellman's gloss of 'exceeds'. His further requirement that commitments be well motivated also has no bearing on the problem outlined in the text. I take it the quotes just inside the parentheses merely signal a *semi*-formal statement, something like the use of quotes in: ZF$\nvdash$ "ZF is consistent", rather than something metalinguistic.

[2]We also leave Fine's third-parameter, if such there be, tacit. (See, chapter 1 (section 2.3).

[3]D_3 is to be the background domain of the context, including anything that can be quantified over by any quantifier in this context.

true in its context of utterance C_1 only if the instance for C_1 is true in C_1; similarly for the other attempts. The following are respectively true in contexts C_1, C_2 and C_3.

(1a) Not everything is quantified over in C_1

(2a) There is a context C' that countenances things which C_2 does not, but not vice versa.

(3a) D_3 does not have everything as a member.

These statements respectively entail the following claims, which are consequently also true in C_1, C_2 and C_3.

(1b) Something is not quantified over in C_1

(2b) Something is not countenanced by C_2

(3b) Something is not a member of D_3

But now the problem is all too apparent. In C_1, the quantifier 'something' ranges only over objects quantified over in C_1. Likewise, in C_2, the quantifier 'something' ranges only over objects countenanced by C_2; and, in C_3, 'something' ranges only over members of the domain D_3 of C_3. The semantics of quantification in these contexts thereby ensures that each attempted statement is true in the original context of utterance only if the corresponding truth-condition holds in the context in which the semantic theory is set forth.

(1c) Something quantified over in C_1 is not quantified over in C_1

(2c) Something countenanced by C_2 is not countenanced by C_2

(3c) Some member of D_3 is not a member of D_3

The truth of (1a) in C_1 leads to an outright contradiction in the metatheory.[4] The first attempt to coherently state relativism therefore faces the following dilemma: either (1) does not imply the instance, (1a), in which case it is too weak to be adequate, leaving

[4]This presentation of the problem broadly follows that of Williamson (2003, pp. 427–8, 430), adapting his arguments to the attempted statements at hand.

open the possibility of the domain of C_1 encompassing absolutely everything; or it does imply this instance, and it is too strong to be true. The same problem faces the second and third attempts.

Hellman, I suspect, will object to the transition from (2a) to (2b). He is well aware of the problem facing extensional claims about the range of quantifiers, and recommends instead an intensional approach to stating relativism. "Instead of futilely saying that there are ('in fact') always more things than can ever be recognised on any given occasion of use or in any context, one can say that, for any context, there is (or can be) another in which the objects *recognised* or *countenanced* go beyond what are recognised or countenanced in the former. What are compared are not asserted in the thesis to actually exist but only to be claimed or taken to exist." (Hellman 2006, p. 77).

It is one thing, however, to be aware of a problem and another to avoid it. And in both his semi-formal statement and his informal exposition just quoted, Hellman—writing, let us suppose in context C_2—uses either the bare plural expression 'things' or the plural definite description 'the things', neither of which is embedded within an intensional context, in the attempt to compare the limited context with the expanded one C'. But neither of these expressions fares any better than plain-old 'something' in transcending the expressive limitations of the context of utterance. Context C' countenances things which C_2 does not, or the things countenanced in C' exceed those countenanced in C_2, only if there is something that C' countenances which C_2 does not. The inference from (2a) to (2b) is sustained, and trouble remains.[5]

The contextual restrictionist should concede, then, that none of these attempts to characterise her view are successful. An exactly analogous problem arises for other forms of relativism about quantification, including interpretational expansionism. The same argument can be run with the more generic notion of perspective replacing that of context. For instance, if the relativist attempts to state her view with (3), she must utter this sentence relative to some perspective p, whether this is construed as a context or an interpretation, or something else altogether. And this faces the same dilemma as

[5]This is not to say that the general strategy must fail. We shall return to the intensional approach in section 4.

before: either the domain of the quantifier 'no domain' encompasses the domain D_p of p, and the statement is too weak, failing to rule out the absoluteness of this domain; or the quantifier does encompass D_p, and it is too strong, implying that not everything in the domain of D_p is in the domain of D_p.

Curiously, the absolutist *can* coherently claim that in the right context (and the right language), the negation of this statement does succeed in stating his view.[6] Suppose that the absolutist utters.

($\neg$3) Some domain has everything as a member
$\exists xx \forall x(x \prec xx)$

When quantification over domains is construed plurally in this way, the context's domain is quantified over in that very context. As such the absolutist can coherently contend that he has said of everything in the absolute domain that it is contained in some one domain.

This will not of course impress the relativist. She does not accept the absolutist's contention that the domain of the context of his utterance has absolutely everything as a member. She maintains therefore that the purport of the absolutist's claim is merely that some limited domain contains everything in that limited domain. And this trivial truth is quite compatible with relativism.

The dialectical situation, then, is far from satisfactory for both sides. When there is a dispute, the normal standard we expect from a statement of the competing positions is for all sides to agree that they succeed in disagreeing: the opponent agrees that the proponent's statement succeeds in stating the proponent's view—even though in his opposition he contends that the view successfully characterised in this way is false. Given this, the dialectic can proceed in the standard fashion: the proponent can support his view, by offering arguments which end with his statement as the conclusion; and the opponent can seek to undermine these arguments or to offer counter arguments, and so on, and so forth.

[6] Compare Williamson (2003, p. 433).

But when, as has so far been the case in the absolutism/relativism debate, the two sides do not agree that they disagree, arguments with a contested statement as the conclusion lose their dialectical force. Consider, for instance, the statement of absolutism, ($\neg$3). For those happy with plural logic—as the absolutist had better be in order to state his view this way—it is all too easy to argue in favour of this statement. One can give an argument in which ($\neg$3) is the conclusion and the premisses are: $\exists xx \forall x(x \prec xx \leftrightarrow x = x)$ and $\forall x(x = x)$. From the perspective of plural logic—where $\forall vv$ is to be read 'for any zero or more things'—the argument is impeccable: it is valid and both premisses are logical truths.[7] But the argument does nothing to advance the absolutist's cause. The relativist can accept the conclusion without conceding a point against her position.

In view of the contextual restrictionist's apparent inability to characterise her view, Williamson (2003, pp. 434–5) proposes that the absolutist opposing her should stake his position on a judiciously chosen context, C. If from the perspective of another context $C^{\star}$, the relativist can show that not everything is quantified over in C, the absolutist should concede defeat. Likewise, if the shift from C to $C^{\star}$ fails to expand the domain, the relativist should concede defeat. The shift from C to $C^{\star}$ is thereby taken to be the 'crucial test' to decide the issue.

But while this gives the absolutist and relativist something to argue about, it falls short of a fully satisfying resolution of their dialectical problems. For, as Williamson observes, 'the link between the result of the test and the underlying issue is not purely logical in either direction' (p. 435). If the crucial test favours relativism, this will embarrass the absolutist, but it does not refute absolutism. For it still remains open that another domain—even that of $C^{\star}$—contains absolutely everything. Likewise, if the test favours absolutism, it is still open to the relativist to attempt a different shift.

Consequently, a full resolution of the statement problem, one that yields formulations of absolutism and relativism that both sides can agree to succeed in characterising the two positions, is desirable for both sides. But the statement problem remains more

[7] Indeed, since the premisses are logical truths, the arguments that result from dropping either or both of the premisses are also valid.

pressing for the relativist. Absolutism appears to be self-reinforcing. From the absolutist's perspective, even a trivial truth of plural logic will succeed in stating his view. But the failed attempts to state relativism suggest that it may be self-defeating. In the examples surveyed, sentences that are sufficiently general to state relativism are false. This also raises worries about whether the relativist can really *believe* the view she purports to believe. If the view cannot be stated in language, why think that it can be captured in thought?

Before I start to set out the principal argument in favour of relativism in part II, therefore, some preliminary work needs to be done to show that there is a tenable position to argue for. The remainder of the present section defends relativism against the most pressing charge: that the view is self-defeating. Section 2 then outlines an account of a schematic form of generality that allows for a characterisation of relativism. This means, together with further expressive resources in the form of suitably interpreted modal operators presented in section 4, is then put to use against two further objections that arise in connection with the limits on quantification posited by relativism. Section 3 presents a simple argument against relativism based on the purported need for absolute generality in metaphysics and logic; section 5 examines an objection concerning more circumscribed 'kind' generality required for science, and ordinary communication, more generally.

1.2 Object- versus metalanguage

The failed attempts of the previous section, together with those elsewhere in the literature,[8] provide inductive evidence that *no* attempt to state relativism will avoid the dilemma that has confronted the surveyed attempts: every attempt will be either too weak to rule out absolutism, or too strong to be true. Setting aside the dialectical problem this causes, the absolutist might be tempted to push the argument further. For—he might argue—the generalised dilemma shows that, for any sentence, either from English, or an interpreted formal language, it is not the case both that the sentence is true and that it rules out absolutism. But, if every true sentence is compatible with abso-

[8]Williamson (2003, §V) surveys a number of further attempts.

lutism, any sentence that states relativism must be false. Relativism—the absolutist concludes—is self-defeating.

I am inclined to accept every point made by the absolutist here, save for the last one.[9] In fact, in addition to the inductive evidence gained from surveying failed attempts, there is a general reason to think that no sentence succeeds in stating relativism. The difficulty is an instance of a more general one: many facts *about* a language cannot be stated *in* the language.

Consider, for example, English*, an impoverished version of English for which the following is the case

> No singular term in English* names Bertrand Russell.

To express this fact,[10] the impoverished language needs a singular term with the same extension as the name 'Bertrand Russell' has in ordinary English. But this is exactly what this fact shows the impoverished language to lack. English* cannot specify what it is unable to name, since to do so it would need to name it.

Attempting to express the fact—assuming it to be a fact—that something lies outside the present perspective's domain faces a similar problem. To state that the domain is not absolute, some sentence, as interpreted relative to this perspective, must name or quantify over an object outside the domain. But it cannot. Every object it names or quantifies over is *ipso facto* a member of the domain.

But the cause of the limitation also points the way to the solution of the statement difficulty. It is clear in general that the expressive limitations of the language being theorised about need not pose a serious obstacle to carrying out such theorising. It is not a problem for the proponent of Russellian anonymity in English* that he cannot state his theory in the language he wishes to theorise about. Instead he states the relevant fact—as we did above—in a suitable metalanguage. Likewise it is no serious obstacle for the relativist who wishes to state that a particular perspective does not have absolutely

[9] Relativists who embrace modal operators of the sort introduced in section 4 will accept the claim that relativism cannot be stated in a language only if the language lacks such modal operators.

[10] We take facts here to be individuated fine-grainedly. A fact is a true structured proposition, built from the semantic contents of the constituents of the sentence that expresses it.

everything is its domain. For although this claim cannot be stated relative to that perspective, the 'object perspective', it can be coherently made from a suitable 'meta-perspective'. For instance, the relativist can unproblematically deny the absoluteness of the domain D_3 of C_3 by uttering (3) relative to a more inclusive perspective.

This, of course, falls short of a general statement of the relativist's theory about generality.[11] But rather than reject it as insufficiently general, the relativist should seek to supplement it. The failure of any *one* true sentence's ruling out absolutism does not show relativism to be false. Nor does it place the relativist in dubious metaphysical or epistemological territory, requiring her to claim that her view is an ineffable truth. It merely shows that relativism is not finitely axiomatisable. Instead of searching for a single truth to characterise her view, circumscribing the limitations of any language, the relativist can take her theory to be made up of many truths, each stating the limitations of a single domain from the perspective of a wider one.

2 Schematic Generality

Relativism, then, is not self-defeating. But it seems that no finite number of sentences will suffice to characterise relativism.[12] This is a not unfamiliar scenario. There are many theories that admit of axiomatisation but not of finite axiomatisation in a given language.

2.1 Open-ended schemas

The standard way to resolve this issue is to employ a schema.

For example, when we take Modus Ponens as the sole rule of inference, we are unable to finitely axiomatise classical propositional logic.[13] Instead we formulate the

[11] Anticipating this sort of move, Williamson (2003, pp. 428–9) raises this objection.

[12] As in n. 9, the relativist who endorses 'inter-perspective' modal operators will only accept this claim for languages lacking the relevant modal operator. See section 4 below.

[13] It *is* possible to axiomatise this theory with only finitely many (non-schematic) axioms in the presence of a uniform substitution rule. As Kleene observes, however, such an axiomatisation must still rely on a rule of inference with 'the character of a schema. . . in order to provide for infinitely many different applications of it' (Kleene 2002, p. 35, n. 28). Von Neumann (1927) is widely credited with replacing a substitution rule with axioms schemas. See, for instance, Quine (1981, p. 89), Church (1996, p. 158), Kleene (2002, p. 35).

axioms schematically. *Inter alia*, the law of contraposition is often formulated with an axiom schema, formulated using the following template:

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

subject to the side-condition that an instance of the schema is obtained by uniformly replacing each occurrence of the schematic variables ϕ and ψ with an object-language sentence.

In general, a schema consists of a *template*, the result of replacing one or more of the expressions in a sentence with schematic variables, and a *side-condition*, which specifies in semantic or syntactic terms which expressions may replace the schematic variables. Any sentence that results from replacing the schematic variables with expressions in accordance with the side-condition is an *instance* of the schema. Infinitely many sentences may thereby be described with a single schema.[14]

Although commonplace, the use of a schema to avoid the charge of being self-defeating may provide new impetus for worries about believability. On the one hand, a theory with infinitely many axioms or instances of rules seems to demand too much of those who uphold it. For how are finite beings like ourselves to be committed to the infinitely many axioms and rules that are the instances of its schemas? On the other hand, at least in the case of logic, the schematic theory seems to deliver too little. As McGee (2000) and Williamson (2006) have each pointed out, our commitment to the *laws* of logic is not limited to upholding the instances of schemas that can be formulated in the present language.[15] It is not as if we uphold the law of contraposition only because we have found English to be sufficiently inexpressive to lack a counterinstance; or as if we find it necessary to renew our commitment when the expressive resources of this language are expanded, checking the law afresh when new expressions enter the lexicon.[16]

[14]See Corcoran (2006) for a historical overview of the use of schemas. We borrow the use of 'template' from his paper.

[15]Williamson makes this point. See Williamson (2006, pp. 376–7).

[16]These points are made by McGee (2000, p. 66) in relation to *reductio ad absurdum*; see also Williamson (2006, p. 376) in relation to disjunctive syllogism.

In view of these apparent problems, we might attempt to eschew a contraposition schema in favour of a single axiom. The infinitely many instances of the schema can be replaced by a single axiom quantifying into sentence position:

$$\forall p \forall q((\neg p \to \neg q) \to (q \to p)).^{17}$$

This may appear to swiftly dispatch both problems. The proponent of propositional logic need only have a single belief, the genuine law of contraposition expressed by the second-order sentence, and this belief underlies every instance of the schema from both the present language and any extension of it.

The trouble with this approach is simple. Eliminating a schema in one part of the theory only causes another to spring up somewhere else. In the higher-order case, in order for example to derive an instance of the propositional schema we need the corresponding instances of an axiom schema of comprehension:

$$\exists p(p \leftrightarrow \phi)$$

or some other schema, or schemas, that yield each instance of this one.

In fact, when one reflects on how proofs are carried out in standard proof systems, it becomes obvious that no finite number of axioms and particular instances of rules, however ingeniously formulated, will yield every instance of the contraposition schema. In a Hilbert-style system we would then have finitely many object-language sentences as (non-schematic) axioms: $A_1, \ldots, A_k$; and finitely many (non-schematic) instances of rules:

$$\frac{\Gamma_1}{C_1}, \ldots, \frac{\Gamma_n}{C_n}$$

where each Γ_j is a set of object-language sentences and each C_j an object-language sentence. Since a proof must terminate either with an axiom or with the application of a rule, the only sentences that can be proved in this system are among: $A_1, \ldots, A_k, C_1, \ldots, C_n$. But these finitely many theorems cannot encompass each of the infinitely many instances

[17] This is in fact a historically retrograde measure, employed, for instance, in Russell (1906).

of the contraposition schema.

The need for a schematic formulation of axioms or rules of inference, then, is not a particular problem for relativism. Sooner or later, everyone—or at least, everyone who wishes to account for our commitment to logic; or by the last argument, *any* system with infinitely many theorems—must accept that this requires commitment to axiom or rule schemas. Fortunately, the attendant epistemic worries, that schemas demand infinitely many beliefs and that these beliefs do not match our commitment, can be met head on.

What is crucial, as McGee and Williamson observe, is that schemas should be considered open-ended rather than relative to the specification of a particular language. Rather than drawing substituends from a fixed lexicon, determining a fixed (albeit typically infinite) set of instances, the template and side-condition continue to determine new instances as the lexicon expands. Commitment to a schema should attach to the schema itself rather than the instances. Rather than taking commitment to a schema to consist in believing each of its infinitely many instances, the commitment consists in a disposition to believe what is expressed by an instance when it is recognised as being an instance of the template whose constituents meet the side-conditions.[18]

2.2 Inter-perspective schemas

Some care is needed, however, with just how schemas are used to delineate relativism. Suppose, for instance, the relativist deploys the following template:

$(R_\varsigma)\ \ \neg\forall x(x \prec mm_\varsigma)$

where ς is treated as a schematic variable.[19]

The plural term that results from replacing ς with a term for a perspective in mm_ς is intended to encode the domain whose non-absoluteness the instance states, and the logical expressions receive their usual interpretation over some *one* domain for every

[18] Williamson (2006, pp. 376–7) makes similar points in relation to open-ended schemas.

[19] A similar schematic statement is proposed by Shaughan Lavine, who proposes that relativism be characterised with the schema $(\exists x)\neg Dx$, with D treated as a schematic variable. He claims that in the case of a predicate encoding the present universe A, '[w]e obtain the claim that something lies outside A, $(\exists x)\neg Ax$, a claim that evidently must belong to a new universe of discourse, broader than A', suggesting in a footnote that the requisite shift might be brought about by a process of accommodation. (2006, p. 142.)

instance. This fails in the now all-too-familiar manner: either the side-constraint rules out substituting a plural term that encodes the domain $\forall x$ ranges over, and no instance rules out this domain's containing absolutely everything; or it does not, and that instance is trivially false. More generally, no schema whose instances are each interpreted over a fixed domain will be able to state relativism, for as in the preceding section, *none* of the instances will be able to refer to or quantify over any object outside this domain.

Instead, as in section 1.2, each instance must be interpreted relative to a perspective with a more inclusive domain than the domain whose limitations the instance states. The relativist needs a schema whose first instances are something like the following.

(R1) $\neg\forall x(x \prec mm_{p_0})$

(R2) $\neg\forall x(x \prec mm_{p_1})$

(R3) $\neg\forall x(x \prec mm_{p_2})$

$\vdots$

The limits of the domain of the first perspective p_0 are stated by (R1) from a second, more liberal perspective p_1, relative to which the plural term mm_{p_0} denotes the members of the first domain and $\forall x$ expresses universal quantification over a second, wider domain. The limits of this second domain are then stated by (R2) from a third perspective p_2, relative to which the plural term mm_{p_1}denotes the members of the second domain and $\forall x$ expresses universal quantification over a third, even wider domain. We then state the limits of the third domain in (R3) from a fourth even wider perspective; and so on.

It is helpful to make the inter-perspective character of such a schema explicit by taking each of its instances to be a labelled sentence, a sentence with a perspective-prefix that indicates the perspective relative to which it is to be interpreted. For example, instances of the labelled schema

$$\varsigma \colon (\neg\phi \ \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \phi)$$

result from replacing the schematic sentence variables ϕ and ψ with sentences, and the schematic perspective letter ς with a perspective term. The resulting instance is a

Table 3.1: The lexicon of $\mathcal{L}^{pl}$

Category e_o	singular variables: $x, y, z, \ldots$
Category ee_o	plural variables: $xx, yy, zz, \ldots$
Category $t/(e_o, e_o)$	the singular identity predicate: $=$
Category $t/(e_o, ee_o)$	the member-'plurality' predicate: $\prec$
Category t/t and $t/(t,t)$	connectives: $\neg$ and $\wedge$
Category t/t	singular quantifiers: $\forall x$, $\forall y$, $\forall z$,. . . .
Category t/t	plural quantifiers: $\forall xx$, $\forall yy$, $\forall zz$,. . . .
Category e_p	singular perspective variables: $r, s, \ldots$
Category e_p	singular perspective constants: $p, q, \ldots$
Category e_p/e_p	the meta-perspective operator: $'$
Category $(ee_o/t)/e_p$	plural term-forming operator: λ

labelled sentence, such as $q_0\colon(\neg P \rightarrow \neg Q_3) \rightarrow (Q_3 \rightarrow P)$. The effect of the instance is to state what is expressed by the sentence $(\neg P \rightarrow \neg Q_3) \rightarrow (Q_3 \rightarrow P)$ as it is interpreted relative to the labelled perspective, q_0.

In general, a labelled sentence is formed by prefixing a perspective term ς to a sentence ϕ to obtain $\varsigma\colon\phi$. Much like before, an inter-perspective schema consists of a template, the result of replacing one or more of the expressions in a labelled sentence, including the label, with schematic variables, and a side-condition, which specifies in semantic or syntactic terms which expressions may replace the schematic variables. Any labelled sentence that results from replacing the schematic variables with expressions in accordance with the side-condition is an *instance* of the schema.

As before, we shall treat such schemas open-endedly. Rather than drawing substituends from a fixed lexicon, the template and side-condition continue to determine new instances as the lexicon, or stock of labels, expands. Commitment to the theory characterised by the schema consists in a disposition to accept what is expressed relative to the labelled perspective by the sentence labelled in the instance, when this labelled sentence is recognised to be an instance of the template whose constituents meet the side-condition.

To formulate relativism we shall deploy a plural lexicon, $\mathcal{L}^{pl}$, with two categories of singular and plural terms: e_o and ee_o for singular and plural object terms, e_p for singular perspective terms. The lexicon is displayed in table 3.1. To allow each meta-perspective p' to describe the domain of the corresponding object perspective p, we equip it with the means to frame a suitable plural term. To this end we shall use *labelled* λ-terms of the form $\langle \lambda_\varsigma v.\phi \rangle$. The intended interpretation of such a term is roughly that of the English term 'those zero or more members of the domain of the perspective labelled by ς that satisfy ϕ relative to this perspective'. With λ-terms thus interpreted, the domain of perspective p is denoted by the plural term $\langle \lambda_p y.\exists x(x = y) \rangle$, which we continue to write as: mm_p.

Relativism may then be axiomatised with the following inter-perspective template, schematising the third attempt to state the view, (3).

Relativism (Schematic Statement). $\varsigma' \colon \neg\forall x(x \prec mm_\varsigma)$

The only schematic element is the schematic perspective letter ς, which occurs both as a constituent in the labelling term ς' and as a constituent in the plural term mm_ς. Provisionally, let us gloss the side-condition as the requirement that the schematic variable be replaced so that ς' denotes a suitable meta-perspective relative to the perspective denoted by ς (we shall return to just how this condition should be characterised below).

The instances of this schema for the perspective terms p_0, p_1 and p_2 are labelled versions of (R1), (R2) and (R3). The labels simply make the perspective relative to which the sentence is to be interpreted explicit. The double occurrence of the schematic variable ensures—given the intended interpretation of $'$—that the sentence in each instance is labelled with a perspective term for an appropriate meta-perspective.

It bears emphasis that the schematic letters in the template are *not* bound by an invisible quantifier, expressing either objectual or substitutional quantification over perspectives. The schema provides a means to describe infinitely many sentences, each of which offers a *partial* characterisation of the view when interpreted relative to the labelled perspective. The schema template itself, in virtue of containing uninterpreted schematic placeholders, does not state anything, let alone relativism. This formulation of relativism does not consist in a single sentence: it is more like a promise to utter certain sentences in certain contexts or under certain interpretations (or relative to certain ontologies), many of which will express propositions that cannot be expressed from the present perspective. *In particular*, instances of the schema jointly tell us, in effect—and *now* I am going to quantify over perspectives—that the sentence that results from replacing the schematic variable ς in the template with the name of *any perspective* is true when interpreted relative to the corresponding meta-perspective. But the schema's instances are not limited to perspective-terms that denote perspectives that we can quantify over in the version of English I am writing in. Commitment to the view whose axiomatisation is described in this way is similarly dispositional: upholding relativism consists in being disposed to believe what is expressed by the sentence labelled in an instance at the perspective indicated by its label, when it is recognised to be an instance.

2.3 The side-condition problem

The generic notion 'perspective' plays a crucial role in the formulation of relativism: the side-condition on the substituends is that ς' should denote a suitable meta-perspective. But it should be borne in mind that different sorts of relativist will characterise this notion of perspective differently. The contextual relativist takes shifts in perspective to consist in shifts in context, the interpretational relativist in shifts of interpretation, and so on.

In the latter case this raises a worry. For in shifting domain even the semantic values of the logical vocabulary such as $\forall v$ is subject to change. Indeed, this is the characteristic claim of the interpretational relativist: shifts in perspective lead $\forall v$ to shifts its interpretation, so that the quantifier comes to express quantification over a wider domain. It is very clear however that not any assignment of semantic value will do. Suppose for instance that, relative to p_1, $\forall v$ expresses *existential* quantification over the domain of this perspective (with $\neg$ and mm_{p_0} receiving their intended interpretations in p_1). Then the first instance of the schema misfires. Rather than stating that not everything is in the domain, the sentence $\neg\forall x(x \prec mm_{p_0})$, as interpreted relative to p_1 with this unintended interpretation of $\forall v$, tells us that *nothing* is in the domain. Clearly for the schematic characterisation of relativism to succeed, the relativist must ensure that the interpretation of each expression that occurs in instances of the schema is strictly constrained.

Here, for a change, the contextual relativist has an advantage over her interpretational counterpart. On her view, shifts in domain are solely due to shifts in context, keeping the interpretation of the lexicon fixed. Consequently she can rely on what she claims to be the standing indexical linguistic meaning of $\forall v$—in Kaplan's (1989) terminology, its fixed character—to ensure that $\forall v$ receives its intended content in each context. The interpretational relativist, on the other hand, needs to articulate a substantive side-condition to accompany the schema template, ruling out unintended interpretations of the quantifiers and other expressions.

The obvious necessary constraint is that the logical expressions should receive their intended interpretations over a more inclusive universe. In particular, $\forall v$ should express universal quantification over a universe wider than the universe M_p of p. To make the requirement precise we shall take it to be partially constitutive of what a perspective is on the interpretational account that the logical vocabulary is interpreted in accordance with the intended Tarskian truth-conditions. An interpretation p' of the lexicon of $\mathcal{L}^{pl}$ is an admissible meta-perspective for p only if the truth-conditions displayed in table 3.2 hold for all assignments σ to singular and plural variables over a universe $M_{p'}$ that is more inclusive than M_p.[20] This partial characterisation of admissible meta-perspective, however, raises a further epistemic concern for the interpretational relativist. The problem is a special case of a more general objection that Williamson (2003, §VI) raises against the relativist deployment of schemas. In order to understand an open-ended schema, the relativist must grasp what distinguishes its instances from its non-instances via a sufficiently general side-condition. The worry is that without absolutely general quantification she lacks the conceptual resources required to frame this condition (pp. 439–40).[21]

[20] In order to make these satisfaction conditions acceptable to the absolutist, assignments to singular and plural variables may be encoded as plural term denotations along the lines of Rayo and Uzquiano (1999, p. 321): an assignment σ is encoded as some pairs of the form $\langle v, a\rangle$ or $\langle vv, b\rangle$ such that for each first order variable v, there is exactly one pair $\langle v, a\rangle$. We write $\sigma(v)$ for the unique a such that $\langle v, a\rangle$ is one of σ and $\sigma(vv)$ for the zero or more objects b such that $\langle vv, b\rangle$ is one of σ. An assignment σ is said to be *over* a plurally encoded domain D if, for each v, $\sigma(v)$ is one of D and, for each vv, each one of $\sigma(vv)$ is one of D.

[21] Williamson raises this problem in the context of discussing relativist attempts to express kind generalisations. See section 5 below. Page references in the remainder of this section are to his 2003 paper.

Table 3.2: Intended semantics for logical expression relative to p'

Let $\mathrm{Den}_\sigma(vv)$ abbreviate 'the (zero or more) objects σ maps vv to'; $\mathrm{Den}_\sigma(\langle\lambda_p v.\phi\rangle)$ abbreviate 'the zero or more members a of M_p such that $\phi[u/v]$ is true relative to p under $\sigma[a/u]$'.

$u = v$ is true relative to p' under σ iff $\sigma(u)$ is identical to $\sigma(v)$, which is a member of $M_{p'}$

$\tau\tau_1 \equiv \tau\tau_2$ is true relative to p' under σ iff $\mathrm{Den}_\sigma(\tau\tau_1)$ are identical to $\mathrm{Den}_\sigma(\tau\tau_2)$, which are members of $M_{p'}$

$u \prec \langle\lambda_p v.\phi\rangle$ is true relative to p' under σ iff $\sigma(u)$ is one of $\mathrm{Den}_\sigma(\langle\lambda_p v.\phi\rangle)$.

$\neg\phi$ is true relative to p' under σ iff ϕ is not true relative to p' under σ.

$\phi_1 \wedge \phi_2$ is true relative to p' under σ iff ϕ_1 is true and ϕ_2 is true relative to p' under σ.

$\forall v\phi$ is true relative to p' under σ iff, for every a in $M_{p'}$, ϕ is true relative to p' under $\sigma[v/a]$.

$\forall vv\phi$ is true relative to p' under σ iff, for all zero or more aa in $M_{p'}$, ϕ is true relative to p' under $\sigma[vv/aa]$.

In the present case, the interpretational relativist's commitment to her view consists, very roughly, in her being disposed to believe that each object language's universe fails to be comprehensive, when she recognises herself to be speaking (or thinking) in an admissible metalanguage, relative to which this claim can be expressed. But to have such a disposition she must have some general grasp of what it takes for a language to be an admissible meta-perspective. Consequently, she needs to frame the side-condition on the schema that characterises her view so that (i) it is semantically adequate, ensuring that instances of the schema template receive their intended truth condition; and (ii) it is conceptually licit, deploying only conceptual resources her view permits. The trouble is that the most obvious means to meet (i) violates (ii).

The obvious proposal is to characterise admissibility directly in semantic terms via the clauses in table 3.2. But, as Williamson observes, semantic conditions of this sort succeed only when the domain of the quantifier $M_{p'}$ is no more inclusive than the universe of the metametalanguage in which the side-condition is framed. The Tarskian clause above is equivalent to this one, where M is the universe of the metametalanguage.

> $\forall v\phi$ is true relative to p' under σ iff, for every a in $M_{p'} \cap M$, ϕ is true relative to p' under $\sigma[v/a]$

But for metalanguage universes that are more extensive than that of the metametalanguage, this characterises universal quantification over the intersection of the meta and metametalanguage universes, $M_{p'}$ and M, rather than over the metalanguage universe. Consequently, unless she deploys conceptually illicit absolutely general quantification, the relativist cannot give a semantically adequate side-condition by setting out the Tarskian clauses relative to any perspective. (Compare pp. 439–40, 444–5.)

The failure of the most obvious characterisation of admissibility, however, does not show that the side-condition problem cannot be met. I would like to suggest a less direct characterisation of admissible meta-perspective. The basic thought behind the style of response I would like to pursue is the following. One major source of information about what our expressions mean are facts about how we and other speakers of the same language use these expressions. And in the case of logical (and mathematical) expressions, one very important feature of use is the patterns of inference that competent language users follow. The proposal, then, is to frame the condition of admissibility in terms of the inferential role that the expressions in the language have. *Inter alia*, for an interpretation p' to be an admissible meta-perspective, the use of logical expressions needs to conform to certain familiar patterns of inference.

Regarding the two desiderata, the approximate thought is that since the meaning of our expressions is determined by our use of them, whether or not a reinterpretation of the lexicon is semantically adequate will be determined by facts about use. Moreover the expressions of such facts does not require expressive resources that the relativist claims to be unavailable. The hypothesis therefore is that use provides a conceptually licit handle on the semantic adequacy of a meta-perspective: a side-condition framed in terms of the inferential behaviour of speakers may meet both desiderata. To examine such a hypothesis will require us to take up some difficult questions in meta-, or foundational, semantics. In particular the question of *how* use determines meaning. As such, we shall not be in a position to fill in the details of this response to the side-condition problem until part III when we take up metasemantic issues is more detail. I am however in a position to respond to an objection to the general strategy.

Williamson claims that no attempt to specify the side-condition is syntactic terms succeeds: 'we can construct a pseudo-quantifier that behaves syntactically but not semantically just like a universal quantifier' (p. 440). Williamson constructs his pseudo-quantifiers as follows. Given a first-order language—for instance, the first-order language of set theory with urelements $\mathcal{L}_{\beta,\in}$—we add 'pseudo-terms' , incorporating 'pseudo constants' $a\hat{}, b\hat{}$ and 'pseudo-variables: $x\hat{}, y\hat{}, \ldots$ and two new pseudo-quantifiers $\forall\hat{}v\hat{}$ and $\forall\hat{}\hat{}v$. Formation of sentences is as usual with $\forall v$ and $\forall\hat{}\hat{}v$ binding only ordinary variables and $\forall\hat{}v\hat{}$ binding only pseudo-variables. Given an interpretation p of the original lexicon, truth-conditions of sentences in the original language are determined in the usual way. But we do not supply extensions for pseudo-quantifiers or pseudo-constants, or assignments to pseudo-variables. Instead, in addition to the usual clauses for connectives, the truth-conditions for mixed sentences—sentence containing at least one pseudo-expression—are settled by the following clauses.[22]

> $P\tau_1, \ldots \tau_k$ is true relative to p under σ iff each term $\tau_1, \ldots \tau_k$ is a pseudo-term.
>
> $\forall\hat{}v\hat{}\phi$ is true relative to p under σ iff ϕ is true relative to p under σ.
>
> $\forall\hat{}\hat{}v\phi$ is true relative to p under σ iff both $\forall\hat{}v\hat{}\phi\hat{}$ and $\forall v\phi$ are true relative to p under σ (where $v\hat{}$ does not occur free in ϕ and $\phi\hat{}$ is the result of substituting each $u\hat{}$ for each free occurrence of u in ϕ).

The semantics for sentences containing pseudo-expressions generates many spurious truths, quite at odds with the intended interpretation that the relativist wishes to secure. Consider for instance the language of first-order set theory with urelements. Suppose that β and $\in$ receive their intended extensions over some inaccessible rank $V_\theta[U]$ based on the urelements in set U. The interpretation of sentences containing pseudo-expressions still goes haywire. Contrary to well-foundedness: $x\hat{} \in x\hat{}$ is true and $x\hat{} \notin x\hat{}$ false under any assignment over the domain. Consequently, $\forall\hat{}x\hat{}(x\hat{} \in x\hat{})$ is true and $\forall\hat{}\hat{}x(x \notin x)$ is false. Clearly, if a side-condition framed in terms of inferential role admits such patently unintended interpretations, it fails to be semantically adequate. But Williamson claims that, with respect to logical consequence, the first pseudo-quantifier $\forall\hat{}v\hat{}$ 'behaves just like a genuine universal quantifier over a domain with exactly one member' (p. 441) and that $\forall\hat{}\hat{}v$ 'behaves. . . just like a universal quantifier over a domain

[22]The clauses in the text modify Williamson's presentation to the conventions in place here, but make no substantial changes. See p. 441

that results from adding one object to the domain of [the original language]' (p. 442). Unless the relativist can explain in terms of use what is wrong with this unintended interpretation, the strategy sketched above for meeting the side-condition problem fails.

Fortunately, there does seem to be a disanalogy between the inferential patterns sustained by ordinary expressions and their pseudo equivalents. In the case of $\forall^{\wedge\wedge}v$, the disanalogy may be found in the classical rule of universal generalisation:

$$\frac{\Gamma \vdash \phi[u/v]}{\Gamma \vdash \forall^{\wedge\wedge} v\phi}$$

(As usual, the rule is subject to the side-condition that u does not occur free in Γ or $\forall^{\wedge\wedge}v\phi$.) Say that a sequent $\Gamma \vdash \phi$ is (materially) *sound* relative to p if, for each assignment σ over the domain of p under which each sentence in Γ is true relative to p under σ, ϕ is also true relative to p under σ. Likewise a rule like universal generalisation is said to be sound if it preserves soundness. When $\forall^{\wedge\wedge}$ expresses universal quantification over the domain of p, universal generalisation is sound but under the unintended semantics for pseudo-expressions it is not sound. For example, given the unintended semantics

$$\frac{\vdash x \notin x}{\vdash \forall^{\wedge\wedge} x(x \notin x)}$$

fails to preserve soundness. This then provides the requisite disanalogy.

In the second case, however, the standard classical rules for $\forall^{\wedge}$ *are* sustained under the unintended interpretation.

$$\forall^{\wedge} v^{\wedge} \phi \vdash \phi[u^{\wedge}/v^{\wedge}] \qquad \frac{\Gamma \vdash \phi[u^{\wedge}/v^{\wedge}]}{\Gamma \vdash \forall^{\wedge} v^{\wedge} \phi}$$

These rules preserve truth under each assignment over the domain and its preservation given that (as is standard for a two-sorted language) we require that $u^{\wedge}$ be a variable of the same sort as $v^{\wedge}$, namely a pseudo-variable. What then is the relativist to say in this case?

The response I would like to suggest is this. The fact that the standard patterns of inference for the universal quantifier are sustained under the unintended semantics does not yet show that there is no disanalogy between the patterns of inference sustained by the two interpretations. The spurious truths and falsehoods are due to the unintended semantics for atomic sentences just as much as the unintended semantics for quantifiers. Clearly we should not expect the intended semantics for the *entire* language to be determined by the patterns of use for quantifiers alone. Now when variables of the form $v^\wedge$ *are* assigned extensions as usual and atomic sentence containing such terms receive their intended truth-conditions, the inferential rules for the quantifier $\forall^\wedge v^\wedge$ above are no longer sound under the *unintended* semantic clause Williamson stipulates for this expression. For example, given a constant $\varnothing^\wedge$ that denotes the empty set, and with $=$ given its intended interpretation over the domain:

$$\forall^\wedge x^\wedge(x^\wedge = \varnothing^\wedge) \vdash y^\wedge = \varnothing^\wedge$$

fails to preserve truth under each assignment. When $x^\wedge$ is assigned to the empty set and $y^\wedge$ is not, the premiss is true and the conclusion false.

We shall return to this problem in part III, where we shall see that given a substantive but independently motivated metasemantic assumption, the right sort of inferential behaviour does determine the intended semantics for the logical expressions of $\mathcal{L}^{pl}$ over a strictly wider universe.[23] For now, however, there are two further worries about expressive power to address.

3 The Simple Argument

Even if the relativist can satisfactorily describe her view, we might still worry that the limitations she posits to address the paradoxes cause trouble elsewhere. In the remaining two sections, I shall argue that this is not the case for either supposedly absolute generalisation, or contra Williamson, for limited generalisation over a single kind of thing.

Start with absolute generality. Here is a simple argument against relativism about quantification, generating an argument against restrictionist or expansionist relativism

[23]See chapter 6 section 3.2.

depending on whether 'absolute' generality is taken to be unrestricted or ultimate generality.

(P1) Some of our utterances attain absolute generality.

(P2) If relativism about quantification is true, then none of our utterances attain absolute generality.

Therefore, (C) relativism about quantification is not true.

In saying that the this argument is simple, I do not say that it is simpleminded. The argument is clearly valid and I think both of its premisses enjoy a good deal of initial plausibility.

This section and the next outline two responses to this argument based on rejecting either the first or the second premiss, opening the possibility of a further bifurcation within relativism.[24]

3.1 Quantificational generality

The most immediate way for the absolutist to support the first premiss of the simple argument is to review some putative examples of quantification over absolutely everything and find them to obviously succeed. Consider again then the mereological nihilist's attempt to state his position with maximal generality. He utters the following English sentence, or its first-order formalisation (relative to a suitable interpretation)

(4) Everything is simple
$\forall x \, \textsc{simple}(x)$

having gone out of his way to bring about a context that is congenial to stating his view in its full intended generality, and explaining that by 'everything' he means 'absolutely everything without any restrictions whatsoever'. A logician might make similar efforts to accompany his statement of the following theorem of predicate logic (with identity) either in English or a suitably interpreted formal language:

[24] Absolutists have tended not explicitly formulate such a head on attack on relativism. The opening pages of Cartwright (1994) perhaps come closest, pp. 1–2.

(5) Everything is self-identical.

$\forall x(x = x)$

The effectiveness of such a direct attack on relativism seems to depend on whether we understand 'absolute' to mean 'unrestricted', as the restrictionist relativist does, or to mean 'ultimate', as the expansionist relativist does. The most natural reading of 'absolute' in the mereologist's elaboration surrounding his statement is as yet another way of pointing to the unrestricted character of his use of the quantifier. The expansionist relativist can agree that *in this sense*—provided the context is right—each utterance *does* achieve absolute—wholly unrestricted—generality. But this does not amount to ultimate generality. To achieve ultimate generality, what the nihilist says must not only tell us that everything is simple; or, semantically ascending, that 'Everything is simple' is true in the version of English spoken by the mereologist; but also, in effect, that this sentence would remain true if the interpretation of the language were admissibly liberalised.

For what it is worth, I do not find it *obvious* that this latter task is achieved in the nihilist's case or *mutatis mutandis* in the case of the logician. To refute the nihilist we need to establish that there is something which is not simple. We certainly do not refute nihilism by coming to speak a new language—say one in which 'simple' means what 'complex' means in the nihilist's language—and then showing that in this new language 'Everything is simple' is not true. But nor is it obvious to me that we would refute the nihilist by showing this sentence not to be true in a language that modifies his language in the more circumscribed way that interpretational expansionists are interested in. The nihilist's characteristic claim does not seem to have implications for whether or not the sentence he uses to make it is true in such languages.

Parallel arguments might be advanced by the contextual restrictionist in response to the corresponding simple argument against her position, with 'absolute' read as 'unrestricted'. For instance, recall that Glanzberg (2006) allows that all local restrictions can be removed so that 'everything' ranges over everything in a contextually determined background domain but maintains that this background domain is also subject to contextual shifts. Now the contextual relativist could exploit this distinction to make a similar argument, claiming that what is obvious is that all local restrictions can be removed, and that what is not obvious is that the domain of our quantifier can encompass everything in every background domain. But whether the point retains anything like its original force when we shift from interpretation to context is doubtful. To return to what would count as a refutation of nihilism: the nihilist ought to concede that, if we can demonstrate the truth of the opposing claim 'Something is not simple' in *any* context, then the view he intended to express fails. Shifts in context are dialectically licit here in a way that shifts of interpretation are clearly not.

While finding it obvious, as many philosophers seem to, that quantifiers can achieve unrestricted generality may convince these philosophers to uphold the first premiss in the argument against contextual restrictionism, it comes too close to a flat rejection of this view to carry all that much argumentative force. But to the extent that this view is widespread, there is another point that can be pressed against the restrictionist in favour of this premiss. If the nihilist, the logician, and many others like them, *do* take their views to be refuted by the falsity of the sentence they uttered in contexts other than the putatively unrestricted one that they went out of their way to bring about, then the restrictionist must judge them to be mistaken about the extent of the generality of their claims. For on his view, the falsity of their sentence in more extensive contexts is compatible with its truth in its actual, restricted context of utterance. The restrictionist is committed to an error theory about our ability to achieve unrestricted quantification.

3.2 'Systematic ambiguity'

To attempt to avoid the charge of attributing widespread error, the restrictionist can look for an alternative means to achieve unrestrictedly general theorising *without* deploying unrestricted quantification. And even if she does not require it to undermine the first premiss of the simple argument, the expansionist might offer analogous means to achieve ultimate generality without deploying ultimate quantification.

Glanzberg (2004) concedes that in some cases we do at least purport to attain unrestricted generality; 'there are, no doubt, a few cases where we really strive to accomplish absolutely unrestricted quantification' (p. 559), citing logical truths as an example. In response, he offers the logician an 'extraordinary measure' with which to ensure that utterances carry 'additional force' (p. 560). That measure is a contextual version of Parsons' 'systematic ambiguity'.

Parsons observes that 'no matter how our conception of 'object' or 'entity' might be expanded, the truth of '$\forall x(x = x)$' will be preserved' (2006, p. 217). In his earlier paper, he argues that while one can take a set-theoretic claim to be interpreted over a particular universe, it could also 'be used to be 'systematically ambiguous' as to what sets the quantifiers ranged over' (1977, p. 512).[25] As Glanzberg puts it, the additional force of such an utterance is that it tells us 'ambiguously, a feature of any domain it might be interpreted as ranging over' (2004, p. 559). Neither Parsons nor Glanzberg advance a detailed theory of just how such 'ambiguity' is to work,[26] but the inter-perspective schemas presented in the last section provide a natural means to regiment this 'extraordinary measure' for contextual and interpretational relativists to achieve absolute generality.

The contextual restrictionist can deploy such schemas as a means to allow the nihilist and logician to characterise their views with the unrestricted generality they strive for.[27] Understood according to the contextual restrictionist's preferred reading, the inter-perspective schema:

(ς: 4) $\varsigma\colon \forall x\, \textsc{simple}(x)$

tells us not only that everything in the present background domain is simple but also that the sentence remains true when we relax contextual restrictions to attain wider background domains.

How successfully schemas provide a means to reject the first premiss, however, is open to question. Schematic generality is an imperfect substitute for quantificational generality. As Williamson emphasises, we cannot embed schemas in larger sentences. Adapting his point to the case at hand, suppose that, moved by restrictionism, the

[25] Indeed Parsons makes the further claim that this is the norm for set-theoretic assertions.

[26] Parsons (2006, p. 217) acknowledges some discomfort at not having more of a theory; Glanzberg (2004, p. 559) restricts himself to a few observations.

[27] Putnam (2000a, pp. 11ff.;2000b, p. 16) and Parsons (2006, pp. 217–8) suggest that systematic ambiguity might have a schematic character, without elaborating on what form these schemas are to take. See also Lavine (2006), who advocates a schematic approach to absolute generality.

nihilist's opponent attempts to characterise her view with unrestricted generality using the following inter-context schema.

(ς:¬4) $\varsigma\colon \neg\forall x\, \textsc{simple}(x)$

Then the anti-nihilist goes too far. The instances of this schema imply that *every* domain contains a non-simple object. In contrast, unrestrictedly general anti-nihilism maintains only that some domain contains such an object. The attempt to negate the schema fails, in effect, by mistakenly giving wide scope to the schematic generality rather than the negation. (See Williamson (2003, p. 438).)

For this reason, schematic generality provides only a limited means of undermining the first premiss. Since schematic generality is always universal, and always takes wide scope, the relativist who takes this line only has a substitute for Π_1-sentences, sentences of the form $\forall \boldsymbol{v}\phi$ where ϕ contains only bounded quantifiers.[28] If the anti-nihilist is able to characterise his view with unrestricted generality, inter-context schemas are not the means by which he does so. To the extent that many theorists take their non Π_1-sentences to achieve unrestricted generality, the restrictionist has yet to advance an alternative to an error theory.

4 Modal Generality

In view of the limitations of schematic generality, the relativist might pursue even greater expressive resources. One means to do so is to introduce an 'inter-perspective' necessity operator $\Box$ into the lexicon.[29]

4.1 Inter-perspective modality

Perspectives are to such a necessity operator much as possible worlds are to the metaphysical necessity operator. To a first approximation, $\Box\psi$ may be glossed as follows.

[28]Linnebo (2006, p. 155) makes a similar point in criticising *typical* ambiguity as a means for achieving cross-category generalisations.

[29]Fine (2006) follows this strategy. As we shall see in part II, Parsons (1983) and Linnebo (2010) also deploy similar modal operators in the context of set theory.

$\Box\psi$ is true relative to p under assignment σ iff ψ is true relative to every perspective q under σ.

Of course the intended interpretation of the inter-perspective necessity operator depends on how perspective is understood: whether a shift in perspective is taken to consist in a shift in context, or interpretation, or in the case of Kit Fine's version of relativism, a shift in the 'third parameter' he postulates. But on all these views, the modality is not circumstantial; it does not concern shifts in world as metaphysical necessity does. Under the metaphysical interpretation of $\Box$—which, to emphasise, is *not* the intended interpretation here—the truth of $\Box\psi$ depends on whether ψ would be true at *non-actual worlds* with ψ receiving its *actual semantic content*, the proposition it expresses in the actual interpretation of the lexicon at the actual context of the utterance. By contrast, when given the intended reading on the interpretational or contextual account, the truth of $\Box\psi$ depends on whether ψ would be true at the *actual world*, with ψ receiving non-actual contents, propositions it expresses in different contexts, or under different admissible interpretations. Or on the third parameter approach, the truth of $\Box\psi$ depends on whether ψ would be true at the actual world and with ψ receiving its actual content, but relative to different values of the third parameter.[30]

The proposed truth-condition for $\Box\psi$ however will be not be adequate by the relativist's lights if it is devoid of modal content. Absolutists and relativists about quantification will disagree about the success of giving the truth-conditions for sentences containing this modal operator in a language that dispenses with it in favour of quantification over perspectives in much the same way that actualists and possibilists about quantification will disagree about the success of giving the truth-conditions for sentences containing the metaphysical necessity operator in a language that dispenses with it in favour of quantification over possible worlds. From the relativist's point of view, the truth-condition for $\Box$ is inadequate because it ignores perspectives that fall outside of the domain of the meta-perspective that the condition is stated in. $\Box\psi$ may be false even though the proposed truth-condition is met when ψ holds relative to every perspective in this domain but fails relative to some perspective outside it.

[30] Compare Fine (2006, pp. 33–5) on postulational possibility.

Rather than attempting to give non-circular truth-conditions for sentences containing this operator, the relativist can elucidate $\Box$ more directly in terms of its inferential behaviour. When Γ is a set of labelled sentences, write $\Gamma \vdash \varsigma\colon \phi$ to indicate that what is expressed by ϕ relative to the perspective labelled by ς may be inferred from what is expressed by the sentences in Γ relative to their labelled perspectives. Then the inferential behaviour of $\Box$ conforms to the following inference rules:

$$\varsigma\colon \Box\psi \vdash \varrho\colon \psi \qquad\qquad \frac{\Gamma \ \vdash \ \upsilon\colon \psi}{\Gamma \ \vdash \ \varsigma\colon \Box\psi}$$

with the right-hand rule subject to the side-constraint that υ be replaced by a perspective variable that does not label any sentences in Γ. The relativist can concede that this inferential behaviour is *analogous* to that of the universal quantifier with perspective variables playing a role *akin* to object variables. The truth of $\forall v\psi$ ensures that ψ is true under any assignment of a member of the universe to v. Similarly the truth of $\Box\psi$ ensures that ψ is true relative to any perspective. And when we can demonstrate that ψ is true under σ, without making any special assumptions about what σ assigns to v (save perhaps that it denotes a member of the domain), we can infer that $\forall v\psi$ is true. Analogously, when we can demonstrate that ψ is true relative to perspective r, without making any special assumptions about r, we can infer that $\Box\psi$ is true relative to p. But this analogy does not show the inter-perspective modal operator to be semantically reducible to quantification any more than the analogy between existential quantification and infinite disjunction shows objectual quantification to be semantically reducible to substitutional quantification.

As usual, the corresponding possibility operator $\Diamond$ may be introduced as an abbreviation for $\neg\Box\neg$. This operator may then be shown to display the following inferential behaviour on the basis of that of $\Box$ and $\neg$.

$$\frac{\Gamma, \upsilon\colon \psi_1 \ \vdash \varsigma\colon \psi_2}{\Gamma, \varrho\colon \Diamond\psi_1 \ \vdash \ \varsigma\colon \psi_2} \qquad\qquad \varsigma\colon \psi \vdash \varrho\colon \Diamond\psi$$

The left-hand rule is subject to the side-constraint that υ be replaced by a perspective variable that does not label, or occur in, any member of Γ, $\varrho\colon\psi_1$ or $\varsigma\colon\psi_2$.

4.2 Expressive issues revisited

The availability of such a modal operator opens the way for the relativist to state what she was previously confined to schematically promising to state. Just as the actualist can simulate the effect of quantification over possibilia by embedding world-bound quantifiers within metaphysical modal operators,[31] so too might the relativist attempt to simulate the purported effect of quantification over absolutely everything by embedding perspective-bound quantifiers within inter-perspective modal operators.

For example, in the case of the nihilist's or logician's utterances, prefixing the original universally quantified sentence within a single modal operator suffices.[32] The relativist who endorses an inter-perspective modal operator can claim that an absolutely general analogue of the two views can be stated with the following modal sentences.

($\Box$4) $\Box\forall x\,\textsc{simple}(x)$

($\Box$5) $\Box\forall x(x = x)$

Given such resources, the relativist has an alternative response to the simple argument to the one outlined in section 3.1. She can accept the first premiss but deny the second. Such a stance amounts to adopting a mixed position on absolute generality. On this view, while quantifiers do not achieve absolute generality, modal operators do. This is particularly welcome for the contextual relativist who has thus far struggled to reject the first premiss of the simple argument. Such operators permit her to reject the second premiss instead. She may contend that by embedding his quantifiers within modal operators, as in ($\Box$4) and ($\Box$5), the mereologist and logician *do* achieve absolute generality, albeit without deploying absolutely general quantification.

So construed, the relativist has no trouble with unasserted instances of unrestricted generality. Since nihilism is now formulated as a single statement, the anti-nihilist has

[31] See Fine (1977, 2003).

[32] For sentence of greater logical complexity, a more complicated translation is required; see the discussion of 'modalisation' in chapter 5 (section 3).

no trouble in stating her opposing view, which may be formulated simply by negating the nihilist's statement: $\neg\Box\forall x\,\textsc{simple}(x)$.[33]

The availability of an inter-perspective modal operator also allows the relativist to replace the schematic characterisation of her view with a modal *statement*. Her theory about generality can now be finitely axiomatised as the modal analogue of (3).

Relativism (Modal Statement). $\Box\forall xx\neg\Box\forall x(\Diamond x \prec xx)$.

Similarly, the absolutist can also state a weak version of his view as the modal analogue of his earlier statement that some domain comprises everything.

Absolutism (Modal Statement). $\Diamond\exists xx\Box\forall x(\Diamond x \prec xx)$.

Typically, however, the absolutist will want to go further and drop the initial $\Diamond$ to claim not merely that we could reach an absolute domain relative to *some* perspective but that we already have reached this domain at the present one.[34]

Such modal resources substantially improve the relativist's ability to engage in absolutely general theorising. But they also promise to significantly improve the dialectical situation not just from the relativist's perspective but also from that of the absolutist. Without such modal operators, neither side could frame a statement that both sides could accept as characterising their position. The absolutist resorted to stating his view as a trivial truth, which only succeeds in stating his view if absolutism holds. With such modal operators, the normal standard expected of view-stating can be met: both sides can frame a statement that succeeds in stating the position whether or not the position is true. Both sides have finally reached a point where they can agree that they have succeeded in disagreeing. Provided, that is, that modal operators with this function are available to us.

[33] Fine (2006, p. 41) suggests just this means for the logician to achieve absolute generality.

[34] Fine (2006) proposes a statement of relativism combining procedural possibility and metalinguistic vocabulary: (setting aside niceties of use and mention, see p. 25) he writes $\exists_I x\phi(x)$ to indicate that there is some x under I for which $\phi(x)$, relativism is then formulated on p. 30 using both quantification over interpretations, and the metalinguistic locutions, and the modal operator as follows (working through definitions).

$$\Box\forall I\Diamond\exists J(\forall_I x\exists_J y(x = y) \wedge \neg\forall_J x\exists_I y(x = y))$$

4.3 The objection from irrelevance

Williamson (2003, pp. 431–3) considers the operator strategy as a means of stating contextual relativism with the English strings 'in any context' and 'in some context' treated as irreducible operators like $\Box$ and $\Diamond$ under the contextual reading, and 'in the actual context' performing the role of an actuality operator, and offers two objections. His lines of objection likewise threaten the above statement without the actuality operator. The first is based on an analogy between metaphysical and contextual modality.

Consider the following consequence of the contextual relativist's modal statement $\Diamond\neg\forall x(x \prec \boldsymbol{mm})$, where $\boldsymbol{mm}$ is claimed by the absolutist to be the absolute domain. Clearly the modal statement of relativism succeeds in stating the view in general only if this claim rules our the absoluteness of the purportedly absolute domain $\boldsymbol{mm}$ in particular.

Williamson argues that statements of this sort are too weak to achieve this via an analogy with the metaphysical case.[35] When $\Box$ is given the unintended reading of metaphysical necessity, Williamson observes that it fails to rule out absolutism: 'Those who regard modal operators as irreducible do not give up unrestricted quantification merely by rejecting the Barcan formula for possibility... In asserting that there *could have been* something other than what there actually is, they do not assert outright that there *is* something other than what there actually is' (p. 432).

So much seems incontrovertible. But this is not yet a problem for the modal statement of contextual relativism, of course, for the *intended* reading of $\Box$ is not circumstantial, and thus neither metaphysical nor temporal. Trouble only threatens when Williamson draws the analogy (treating 'in some context' and 'in the current context', recall, as operators, not quantifier expressions): 'Analogously, those who regard contextual operators as irreducible do not give up unrestricted quantification merely by rejecting the Barcan formula for contextuality.... In asserting that, in some context, there is something other than what, in the current context, there is, they do not assert outright that there is something other than what, in the current context, there is.'

[35] Williamson's criticism focuses on a statement deploying actuality operators instead of plural terms but the differences do not affect the wider point.

(p. 432)

Now if this is trouble for the context-operator statement, we can expect a similar problem to arise for the modal statement under the interpretational and postulational readings. However there appears to be no real trouble in the contextual case. The analogy breaks down at the crucial point. Unlike expansion through a change of circumstances, expanding the range of purportedly unrestricted general quantifiers by changing the context *does* refute their unrestricted generality. Unlike in the temporal case, the relativist *does* contradict the unrestrictedness of $\boldsymbol{mm}$ if (without changing the circumstances) she can succeed in shifting to a more extensive context and truly utters $\neg\forall x(x \prec \boldsymbol{mm})$ in that context. Indeed, in his proposal for the terms of the absolutism/relativism debate, Williamson suggests exactly this sort of shift as the 'crucial test' on which both the absolutist and relativist should stake their position. (2003, p. 435). But if, in any context, such contextually-specific claim rules out the unrestrictedness of $\boldsymbol{mm}$, so too does a contextually-general modal claim $\Diamond\neg\forall x(x \prec \boldsymbol{mm})$ that such specific claims bear witness to. For the inferential behaviour of $\Diamond$ is analogous to that of $\exists$. If we can infer that absolutism fails on the assumption that $\neg\forall x(x \prec \boldsymbol{mm})$ holds relative to r—without making any further assumptions about r; then we can likewise infer that absolutism fails on the assumption that $\Diamond\neg\forall x(x \prec \boldsymbol{mm})$ holds relative to the perspective at which it is uttered.

4.4 The objection from 'monsters'

Williamson's second objection stems from the context operator being, in David Kaplan's expression, a 'monster'.[36] Williamson observes that, in English, operator-type expressions like 'in some context it is the case that ϕ' do not perform the same role as metalinguistic expressions like 'ϕ is true in some context'. To borrow his example, suppose someone other than Tony Blair utters the following sentences.

(6) 'I am Tony Blair' is true in some contexts.

(7) In some context it is true that I am Tony Blair.

[36] Kaplan (1989, p. 510ff.).

The former, metalinguistic claim is true. In contexts where Tony Blair is the speaker, the extension of 'I' is Tony Blair. But the English expression 'In some context it is the case that…' does not effect similar shifts of content. In the context of the second utterance, 'I' refers to the speaker. The sentence says falsely of someone who is not Tony Blair that in some contexts he or she is. (See (2003, pp. 432–3).)

Analogous problems arise for attempts to render the interpretational operators in English as 'In some language it is true that…'.

(8) 'Grass is red' is true in some languages.

(9) In some languages it is true that grass is red.

The metalinguistic claim is true, witnessed for instance by languages like English save that 'red' means what 'green' means in English; but the latter claim does not have the same effect in shifting the interpretation of 'red'.

So much seems clear but whether or not the point is damaging depends on what use the relativist wishes to put such operators to. The 'monsters' point casts doubt on any attempt to claim that we achieve absolute generality by deploying such operators *in English*. But such observations about 'monsters' in natural language have no bearing on whether such operators might be introduced into an artificial language, a possibility which Kaplan is explicit in leaving open. (1989, p. 510.) Consequently, when it comes to theorising that requires absolute generality, the relativist is in a similar position to the absolutist who wishes to give an extensional or categorial semantic theory for a language whose universe he claims to be absolute. Given that English lacks superplural quantification and monstrous operators, neither the absolutist nor the relativist can state their theory in natural language. Each must therefore rely on formal languages that transcend the expressive capabilities of English and may therefore need to be acquired, in Williamson's phrase, by the 'direct method' (2003, p. 459).

In the case of semantics, this came at the cost of committing the absolutist to an error theory about empirical semantics. Contrary to widely accepted linguistic practice, the absolutist must deny that the extensional or categorial semantics of sufficiently rich fragments of English can be carried out in natural language (or first-order set theory).

The contextual relativist faces a similar problem. To the extent that English utterances are widely taken to achieve unrestricted generality, in the absence of a replacement for context operators, the contextual restrictionist is also committed to an error theory. But, if the argument above in section 3.1 is right, the interpretational expansionist faces no similar problem. There appears to be no similar presupposition of the availability of ultimate generality in English.

4.5 Thorough-going relativism versus pluralism

The introduction of inter-perspective modal operators leads to a further schism within relativism, a division that cuts across both the restrictionism/expansionism and the interpretational/contextual divides.

Faced with the simple argument the relativist has two choices. The most obvious response is to deny the first premiss, and contend we never attain absolute generality, no matter what expressive resources we might draw. Let us call this position thorough-going relativism. The other response is to adopt a mixed view. Contrary to her through-going counterpart, this sort of relativist accepts the first premiss of the simple argument, conceding that some of our utterances do achieve absolute generality, but denies that such generality is achieved by quantifiers alone. Such a relativist may contend for example that while quantifiers are unable to range over absolutely everything, we can nonetheless achieve absolute generality by embedding our quantifiers within inter-perspective modal operators. Let us call this position pluralism about generality: the proponent of this view is, in effect, a relativist about quantifiers but an absolutist about modalised quantifiers.

The thorough-going relativist need not deny that pluralist's inter-perspective modal operator is intelligible, any more than the relativist need deny that quantifiers are intelligible. But to take a thorough-going view she must deny that either attains absolute generality. Just as no quantifier is able to range over absolutely every object, the thorough-going relativist claims that no inter-perspective operator ranges over absolutely every perspective. No matter how extensively $\Box$ may be interpreted, we can always come to more inclusive understandings of the modal operator.

The pluralist's limited acceptance of absolute generality gives her an expressive advantage over her more austere counterpart. For instance, while the thorough-going relativist must content herself with a schematic characterisation of her view, the pluralist is able to actually state it. But the acceptance of absolute generality also leads to a worry for this mixed view. The worry is that that the sort of motivations that lead to relativism about quantifier might also motivate relativism about modal operators. The pluralist may have to give up on some on the distinctive advantages that led her to repudiate absolutism in the first place. We shall return to further assess the relative merits of through-going relativism versus pluralism in parts II and III. For now there is one further expressive issue to address.

5 The Kind Generality Problem

The final problem pressed against relativism that we shall consider concerns generalisations that Williamson labels 'kind generalisations'. Whether or not we attain absolute generality with our quantifiers, it is obvious that we both have and need a more circumscribed form of generality, the ability to generalise, for instance, about every set of real numbers, or electron, or donkey. We need such bounded generality in mathematics, science and ordinary talk. And in the right context we achieve it when we utter sentences like the following.[37]

(10) Every bounded set of real numbers has a least upper bound

(11) Every electron moves at less than the speed of light

(12) No donkey talks

Williamson argues however that even such limited generality is beyond the reach of the generality relativist. If this were correct, the objection would severely reduce the appeal of relativism. For this view would render what is manifestly expressible inexpressible, and severely curtail our ability to make generalisations that are commonplace in science and mathematics. It will be the business of this section to argue on the contrary that relativism does no such thing.

[37] The last two examples are from Williamson (2003, p. 436).

5.1 Inextensible kind generalisations

Williamson glosses kind generalisations as 'uncontroversially intelligible universal generalizations about absolutely every member of a comparatively limited kind' (Williamson 2003, p. 436). Such generality is not always achieved by restricted quantifiers. For it requires generality over every member of the relevant kind. We do not achieve the relevant sort of generality if we utter (10) in a context in which 'every bounded set of real numbers' is restricted only to sets of real numbers definable in a given language, or if we utter (12) in a context in which 'no donkey' is restricted to donkeys on Blackpool beach.

But nor can any putative generalisations over absolutely every member of a comparatively limited kind count as a kind generalisation if they are to be uncontroversially intelligible. The relativist will contend that while we can quantify without further restriction over absolutely every donkey, or the like, quantification that purports to be over absolutely every member of an indefinitely extensible kind, such as 'set' (*simpliciter*), 'ordinal', or 'interpretation', fails. And we have already reviewed the relativist options in response to this failure in the previous section. Kind generalisations, then, must concern inextensible kinds.

5.2 The regress objection

On the face of it kind generalisations do not appear to require absolute generality. The third example is representative. The obvious way for the relativist to attempt to make the intended kind generalisation about absolutely every donkey is to utter (12) relative to a perspective whose domain contains every donkey, a perspective at which the condition 'donkey' is inextensible. And there seems every reason to think this succeeds.[38] Given that no further donkeys are added to the domains of more extensive perspectives, quantification over this perspective's domain ranges over absolutely every donkey. Williamson (2003, §VI) argues, however, that even quantification over a domain containing every member of the kind does not suffice to make a kind generalisation.

The objection comes in two parts. The first part consists in the following datum:

[38] As Parsons puts it, 'Why should donkeys be worse off than, say, sets of finite rank? We know where counterexamples to 'No donkey talks' are to be found if they exist. . . ' (2006, p. 218).

'Whether the domain contains absolutely every member of the relevant kind is a purely extensional matter; what one says by uttering a generalization over a domain is an intensional matter' (p. 437). Williamson supports this datum with two examples, the second of which is representative. (To gain flexibility for the relativist, Williamson is here supposing that she works in a formal language in which different quantifiers: $\forall[1]$, $\forall[2],\ldots$, are attached to different domains.)

> To say that absolutely no donkey talks, it is not sufficient to utter the formula $\forall[1]x(Dx \supset \sim Tx)$ with the intended interpretation of the predicates over a domain that in fact contains absolutely every donkey. For example, one might stipulate that the domain is to contain all and only terrestrial animals under the mistaken impression that there are talking donkeys on Mars. By uttering that formula with the interpretation, one would not say that (absolutely) no donkey talks, even though the domain does in fact contain absolutely every donkey, since it contains absolutely every donkey on earth and in fact absolutely every donkey is on earth (we may assume). (p. 437.)

The second part of the objection consists in a vicious regress argument that follows directly after his presentation of this example. Williamson continues:

> To say that absolutely no donkey talks, speakers need some sort of access to the information that absolutely every donkey is in the domain. To express that information, we need a formula such as $\forall x(Dx \supset \exists[1]y\, y = x)$ ('Absolutely every donkey is identical with something in the first domain'). But in the current language the outer quantifier needs a domain index too, as in $\forall[2]x(Dx \supset \exists[1]y\, x = y)$. But to express the information that absolutely every donkey is in the first domain we must *already* have access to the information that absolutely every donkey is in the second domain. This looks dangerously like the start of a vicious regress: to have access to one piece of information in the series, we need *prior* access to a previous piece of information in the series. (p. 437, my emphasis.)

Let us start with the second part. We should distinguish two versions of the access claim that is the engine behind this regress. The stronger version adds the word 'prior' to the weaker claim that generalises Williamson's formulation in the obvious way.

The Access Thesis. To generalise over every member of an inextensible kind, speakers need some sort of access to the information that absolutely every member of that kind is in the domain.

The Prior Access Thesis. To generalise over every member of an inextensible kind, speakers need some sort of *prior* access to the information that absolutely every member of that kind is in the domain

Now even the weaker thesis seems doubtful to me. We need not have perfect internal access to the external meanings of our words, especially if these are determined in part by extra-linguistic facts or speakers other than ourselves.[39] But whether or not a quantified sentence permits us to express a kind generalisation depends on the meaning of the quantifier When factors beyond our epistemic ken contribute to ensuring that the quantifier ranges over absolutely every member of a kind, might not this push knowledge of whether or not the quantifier has the right sort of meaning out of reach?

Setting this point to one side, neither of these theses affords either party in the absolutism/relativism debate an advantage over the other. The first is equally harmless to both sides; the second, equally harmful. In the first case, the Access Thesis lacks the antecedence requirement needed to get a regress going. Take the absolutist first. To generalise over absolutely every donkey when he utters $\forall x(Dx \to Tx)$ (under its intended interpretation) he must, by the Access Thesis, have access to the information that every donkey is in the domain of $\forall x$. But this further information is readily available, expressed by $\forall x(Dx \to \exists y\, y = x)$.[40] Without an antecedence requirement, there is nothing to stop a sentence being the route to the information required for that very sentence to express a kind generalisation. The second sentence thereby provides the access to the information that the Access Thesis maintains is required for *both claims* to achieve the requisite generality. But the relativist can make an exactly parallel argument, replacing $\forall$ with $\forall[1]$ interpreted over a suitable domain—in particular, a domain that includes absolutely every donkey. As before, $\forall[1]x(Dx \to \exists[1]y\, y = x)$ succeeds in saying that absolutely every donkey is in the domain of $\forall[1]x$ in the first sentence, thereby providing the access required by the Access Thesis for *both claims* to achieve the requisite generality.

[39] Consider, for instance, Putnam (1981, ch. 1) and Burge (1979).

[40] This latter claim is, of course, a logical truth, but since—we may suppose—a suitable domain has be supplied for $\forall x$, in particular, a domain that includes absolutely every donkey, this still succeeds in saying, in effect, that absolutely every donkey is identical with something in the domain of $\forall x$ in the first sentence.

The second, stronger thesis generates trouble for both sides. According to the Prior Access Thesis, for the absolutist to generalise over absolutely every donkey with $\forall x(Dx \to Tx)$, he must have *prior* access to the information that absolutely every donkey is in the domain of $\forall x$. (Call this information I.) The absolutist has prior access to the information I only if there was a first occasion—at t_0, say—on which he had this access. To access the information at t_0 he needs to use a sentence like the one he used last time, $\forall x(Dx \to \exists y\, y = x)$. Applying the Prior Access Thesis once more, this sentence will not succeed in generalising over absolutely every donkey and provide access to the information I unless he also has prior access to the information that absolutely every donkey is in the domain of the quantifier in this sentence, $\forall x$; in other words, unless he has access to I prior to t_0. But, he does *not*: t_0 is the first occasion of access; and so the absolutist fails to generalise over absolutely every donkey with the first sentence. The relativist faces similar trouble.

Consequently, if the second part of this argument is to be a dialectically effective weapon for the absolutist to wield against relativism, Williamson needs to establish two things. First, that the relativist is committed to the stronger Prior Access Thesis. Second, that the absolutist is not so committed.

Take the second point first. Williamson writes: 'Generality-absolutism faces no corresponding problem of principle, for it treats the presence or absence of contextual restrictions as a feature of the relevant semantic structures. When one omits all such restrictions, one makes an absolutely unrestricted generalization by default' (p. 437). But an appeal to semantic considerations cannot succeed in cleaving a difference between absolutism and *expansionist* relativism. For the expansionist relativist can agree with the absolutist on the semantics of quantification. (See chapter 2.)

On the first point, the considerations Williamson raises here do not demonstrate the relativist's commitment to the Prior Access Thesis. In particular, the datum that immediately precedes Williamson's formulation of the Access Thesis is readily explained without appeal to this thesis or the regress-causing one. Williamson is quite right about the datum: if the quantifier $\forall[1]x$ in $\forall[1]x(Dx \to Tx)$ is restricted to terrestrial animals, then in uttering this sentence one does not say that (absolutely) no donkey talks. But this failure can be readily explained without appeal to problematic theses about access. It fails to express the kind generalisation because it delivers the wrong truth-conditions. In a possible world w in which there are talking donkeys on Mars but none on Earth, the kind generalisation is false but what is expressed by the sentence uttered in the restricted context is true, since some donkey talks in w but no *terrestrial* donkey does.

More generally, Williamson is correct to observe that it does not suffice to express a kind generalisation about absolutely every donkey that the domain in fact contains absolutely every donkey. But the reason for this is that it is a necessary condition for making such a generalisation that *necessarily* the domain contains absolutely every donkey. Domains that contain absolutely every donkey in this world but fail to in others deliver the wrong truth-conditions, by ignoring counterfactual counter-instances to the kind generalisation. This however is no more a problem for the relativist than for the absolutist. Explicit and contextual restrictions must be intensional in character. The restriction on 'every donkey' is not to the set of donkeys whose membership is invariant across the worlds at which it exists but to the condition *donkey* which applies at each world to the donkeys in that world. Provided that in each world, any wider constraint imposed by the relativist—the background domain determined by the determiner, or the universe of the entire lexicon—also encompasses every donkey, there is no barrier to (12) receiving the intended truth-conditions.

Part II

Mathematics

4 Arguments from Paradox

1 Introduction

Recall the informal argument against absolutism based on the reasoning of Russell's paradox. Given any sets—in particular given the sets that the absolutist claims to comprise absolutely every set—their 'Russell set', cannot be one of them.

> Consider those sets amongst the given ones that lack themselves as elements. Now consider the Russell set, which has these non-self-membered sets as its elements. Suppose for *reductio ad absurdum* that the Russell set *is* amongst the sets we started with. If the Russell set is non-self-membered, it is one of the given sets that is non-self membered, and hence one of the elements of the Russell set. If it is self-membered, since the Russell set has as elements only sets that are *non*-self-membered, it is not an element of the Russell set after all. This yields a contradiction since it *is* the Russell set. The Russell set is self-membered just in case it is not.

The principal aim of this chapter is to elucidate and regiment this style of argument and assess its effectiveness against absolutism.

The thought that the set-theoretic paradoxes threaten quantification over all sets is an old one. Within a few years of his letter to Frege, Russell sought to avoid the paradoxes by denouncing quantification over all sets as meaningless. The case was famously taken up by Dummett, and has been much revisited since.[1] All the same, to

[1]See Russell (1906), Dummett (1983, chs. 15–6; 1991, ch. 24). More recently the paradoxes have been exploited to object to absolutism by Fine (2006), Glanzberg (2004, 2006), Hellman (2006), Shapiro and Wright (2006). Boolos (1993), Cartwright (1994) and Williamson (2003) defend the absolutist.

the best of my knowledge, no treatment of the paradoxes has yet sought to explicitly integrate the best known argument for relativism about quantification with a general characterisation of this view.[2] Ordinarily we might expect an argument for a position to end with the statement of the view as its conclusion. But in the wake of the statement problem, arguments for relativism have been hampered by the unavailability of a general formulation of this theory. In the previous chapter, however, we saw that the statement problem can be met. Given either inter-perspective schemas or modal operators, the relativist can succeed in characterising her view. In this chapter, we will take up the question of how the relativist can *reason* with such resources, and on this basis offer regimentations of the argument from paradox, from highly plausible premisses to a suitable formulation of this view as the conclusion.

In view of the amount of discussion the argument from paradox has received, however, it will be as well to start by distinguishing the line of argument I wish to pursue from some less effective arguments in its vicinity.

One style of argument against absolutism that I have already pledged to avoid uses the All-in-One Principle, the assumption that quantification presupposes a set- or set-like domain to quantify over. One lesson of the Russell argument, applied within standard set theory, is that there is no universal set V, with every set as an element. Given All-in-One, it follows that no quantifier ranges over every set. Thus, no quantifier is absolute.[3] But, as we saw in the introduction, the absolutist is under no pressure to accept the key assumption. The absolutist may simply take the argument to constitute a refutation of the All-in-One principle.

Another style of argument that I shall not pursue is one that relies on the Naïve Comprehension schema. It may be thought that the informal argument as given above, and the 'intuitive' conception of set that underlies it, are nothing other than the Naïve

[2]Fine (2006) comes closest. He offers a modal formulation of relativism (see ch 3) and criticises non-modal arguments for this view for being insufficiently general. But he does not formulate a modal argument to replace the non-modal one, and writes: 'it is not my *principal* concern to argue for that position but to show that there is indeed a position to argue for.' (His emphasis, p. 30).

[3]Cartwright (1994, p. 17) accuses Dummett of implicitly relying on All-in One in this way. See also Boolos (1993, p. 222).

conception.[4] The crucial step in the argument is to suppose that *there is* a Russell set for the given sets. But standard Zermelo-Fraenkel set theory tells us no such thing. Instead the argument—this line of thought continues—must rely on the Naïve Comprehension schema to deliver the requisite set.[5]

Naïve Comprehension. $\exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow \phi(z))$

So construed the argument is clearly hopeless. For another lesson of the Russell paradox is that the instance of this schema for $\text{ß}z \wedge \neg z \in z$ is refutable in classical first-order logic,[6] and indeed in weaker systems, including intuitionist and free logic. But from an inconsistent premiss, we can reach *any* conclusion. We could just as easily prove absolutism, rather than relativism. An argument with an inconsistent premiss provides no support for its conclusion.

This regimentation of the argument, however, is not especially faithful to the informal argument. For one thing, the informal argument obtains the Russell set in two steps, not one. First we obtain, severally, those sets amongst the given ones that lack themselves as elements, and then from these sets, we obtain the Russell set. We can likewise replace the Naïve Comprehension schema with two principles. Let us introduce a predicate to say that some objects comprise only sets: $\text{ß}vv =_{\mathrm{df}} \forall x(x \prec vv \rightarrow \text{ß}x)$. Then these principles may be stated as follows:[7]

Plural Comprehension. $\exists xx \forall z(z \prec xx \leftrightarrow \phi(z))$

Collapse. $\forall xx(\text{ß}xx \rightarrow \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec xx)))$

(with Plural Comprehension subject to the side condition that $\phi(z)$ not contain xx free.)

As in earlier chapters, the intended reading of $\exists vv$ is ‘some zero or more things’.[8] Plural Comprehension instructs us to be maximally liberal about which conditions de-

[4]See for instance Uzquiano (2009, p. 306).

[5]Often this principle is formulated with the set predicate ß. Its inclusion will emerge below as something of a concession to the absolutist.

[6]This point is emphasised by Cartwright (1994, pp. 11–2) (in relation to the ß-free formulation of this principle).

[7]The label ‘Collapse’ is from Linnebo (2010) who discusses such a factoring. Linnebo’s formulation is slightly stronger: omitting the restriction to ßv and ßvv. See also Uzquiano (2009) for discussion.

[8]This reading of the plural quantifier in the context of plural set theory has a precedent in Burgess and Rosen (1997), Burgess (2004) and Linnebo (2010) and greatly simplifies the

termine some zero or more things. This principle ensures that there are zero or more non-self-membered sets, severally (without assuming that there is a non-self-membered set in the domain).[9] Collapse instructs us to be maximally liberal about which sets form a set. Given the non-self-membered sets, Collapse ensures that the Russell set exists.[10]

Consequently, a less flat-footed argument against absolutism might seek to blame Plural Comprehension for the paradox, and use this as a means to attack absolutism. For example, much as Naïve Comprehension can be split into the Separation axiom schema and the existence of a universal set, Plural Comprehension can be similarly factored into a separation-like comprehension principle and the thesis that some domain is comprehensive, or all-inclusive:

Plural Separation $\forall yy \exists xx \forall z(z \prec xx \leftrightarrow \phi(z) \wedge z \prec yy)$

Comprehensive Domain $\exists xx \forall z(z \prec xx)$

(with Plural Separation subject to the side condition that $\phi(z)$ not contain xx or yy free.)

The three principles, Comprehensive Domain, Plural Separation and Collapse are individually, and indeed pairwise, consistent in first-order logic supplemented with the analogues of the classical quantification principles for plural quantifiers, but jointly inconsistent in this system.[11] Comprehensive Domain and Plural Separation lead to Plu-

presentation of plural set theory. But nothing substantive hangs on it. Anyone worried about the import of dubious metaphysics in the form of 'the empty plurality' can adapt Boolos's (1984) trick to translate sentences using the 'zero or more' ($\forall_{\geqslant 0}$) reading in terms of the 'one or more' reading ($\forall_{\geqslant 1}$) as follows:

$$\mathrm{Tr}(\forall_{\geqslant 0} vv\phi) =_{\mathrm{df}} \forall_{\geqslant 1} vv\phi \wedge \phi[u \neq u/u \prec vv]$$

where $\phi[u \neq u/u \prec vv]$ is the result of substituting each atomic sentence of the form $u \prec vv$ with a contradiction $u \neq u$. Moreover Burgess and Rosen (1997, pp. 151–5) show how this trick may be extended to interpret the 'zero or more' reading with a 'two or more' reading. The arguments from paradox could then be run *mutatis mutandis* with the translations of the premisses deployed in the text.

[9] On the 'one or more' reading we would need to restrict Plural Comprehension to non-empty conditions:

$$\exists z\phi(z) \rightarrow \exists_{\geqslant 1} xx \forall z(z \prec xx \leftrightarrow \phi(z)).$$

This is equivalent to Tr(*Plural Comprehension*), under the translation defined in n. 8.

[10] Compare Burgess (2005, p. 47) on the source of the contradiction in Frege's logical system.

[11] Consistency may be established by exhibiting Henkin models satisfying each of the pairs of axioms. As usual, a Henkin model is of the form: $\langle D, D^{pl}, |{=}|, |{\prec}|, |\beta|, |{\in}| \rangle$ where D and

ral Comprehension, and with Collapse, to Naïve Comprehension and a contradiction. Consequently we can reach the negation of Comprehensive Domain, by *reductio*, from premisses that do not themselves lead to a contradiction in first-order logic enriched with principles governing plural quantification.[12]

This however is still *not* the argument I wish to pursue. The principal problems with attempts to restrict plural comprehension are two. In the first place, as we shall see in the next chapter, they have tended to lead to a severe mathematical cost, leaving us unable to recover anything like what we expect from set theory. But the main reason I will not pursue this line of argument is that it gives up a principle I find myself unable to disbelieve.

Under the intended reading of the plural quantifier $\forall vv$, each instance of Plural Comprehension is obviously true: given any condition ϕ, there are some zero or more things—namely, the ϕs—that comprise every ϕ and nothing else. Accordingly, my discussion of the arguments from paradox will take place against the background of plural logics which extend the logic governing the first-order fragment of the language not only with plural analogues of the axioms and rules governing singular quantifiers, but also

---

D^{pl} are non-empty sets; $|{=}|, |{\in}| \subseteq D \times D$; $|{\prec}| \subseteq D \times D^{pl}$ and $|\beta| \subseteq D$. Truth in such a model is defined via the usual satisfaction conditions displayed in table 4.2, on page 112, below (replacing 'relative to p' with 'in the model', and 'in D_p' and 'over D_p with 'in D' and 'in D^{pl}).

The pairs of axioms are then true in the following models. Say that a model is *ordinary* if $|{=}| = \{\langle a, a\rangle \in D \times D\}$, $|{\prec}| = \{\langle a, A\rangle \in D \times D^{pl} \mid a \in A\}$, $|\beta| = D$ and $|{\in}| = \{\langle a, b\rangle \in D \times D \mid a \in b\}$.

Plural Separation + Comprehensive Domain hold in any ordinary, full model; that is, any ordinary model with $D^{pl} = \wp(D)$.

Plural Separation + Collapse hold in any ordinary model where D is a supertransitive set—a set that has as elements both the elements and subsets of its elements—and $D^{pl} = D$. For example, taking D and D^{pl} to comprise the hereditarily finite pure sets provides a model.

Collapse + Comprehensive Domain hold for example in the non-ordinary model such that $D = D^{pl} = |\beta| = \{\varnothing\}$; and $|{=}| = |{\prec}| = |{\in}| = \{\langle\varnothing, \varnothing\rangle\}$

It follows that the pairs of axioms do not prove any classical propositional contradiction $\perp$ in the plural logic consisting of the axioms and rules of FPFO less PC and the rules for the plural identity predicate. It is readily verified (i) that in each model such that $|{=}| = \{\langle a, a\rangle \in D \times D\}$ the axioms of this system are true in the model under each assignment of singular and plural variables to members of D and D^{pl} respectively and the rules of this system preserve truth-in-the-model-under-each-assignment (ii) any propositional contradiction $\perp$ is false in each model under each assignment.

[12] Uzquiano (2009) critically discusses two similar arguments, based on Collapse against Comprehensive Domain. The one most similar to the one in the text replaces Plural Separation with the following schema: $\forall y \exists xx \forall z(z \prec xx \leftrightarrow z \in y \wedge \phi)$.

with Plural Comprehension.[13] By this standard, arguments against absolutism based on Collapse fare no better than those based on Naïve Comprehension. Collapse *on its own* is inconsistent in plural logic.[14] That it still leads to an inconsistency in combination with absolutism fails to distinguish this view from any other theory.

Of course, once Plural Comprehension is granted as a logical principle, the Comprehensive Domain thesis emerges as a logical truth, and any premisses that are inconsistent with it are thereby inconsistent *simpliciter*. But, as we saw in chapter 3, the relativist is under no pressure to accept this as an adequate characterisation of absolutism. Rather than formulating his position as a trivial truth, the absolutist should have enough strength in his convictions to characterise a position that the relativist disagrees with. He should contend not merely that some things comprise everything in the present domain, but that these same things continue to comprise everything no matter how the relativist may shift the perspective. Accordingly, the relativist need not maintain—as Collapse impossibly contends—that some set in the domain of the initial perspective comprises every non-self-membered set in that domain. It is enough to demonstrate merely that there is such a set in the domain of a more liberal perspective.

The regimentations of the argument from paradox that I shall present weaken Collapse by making an allowance for such a shift in perspective. I shall present three versions of the argument. Each can be seen as a regimentation of the informal argument set out above; and each in some respects generalises the last; but they also differ as to the ideological resources used to frame them. The first is set out in a plural language; the second and third respectively draw on the schematic and modal resources introduced in the previous chapter. The next section formulates the arguments, and identifies and defends the underlying logic each deploys. The following two sections then defend their principal premisses.

[13] Or, in languages with λ-terms, λ-exchange principles.

[14] This is not to say that other principles *analogous* to Collapse must suffer a similar deficiency. In particular, as we shall see, this argument does not apply to a modal analogue of this principle.

2 The Arguments from Paradox

2.1 The metalinguistic argument

As a warm-up case, let us start with an argument that is not very ideologically demanding but also rather less general than the relativist should aim for. The argument is set out in a plural language $\mathcal{L}_1^{pl}$, an extension of $\mathcal{L}_{\text{ß},\in}$, which enriches the first-order language of set theory *inter alia* with plural variables and quantifiers, the member-'plurality' predicate $\prec$ and the 'plurality' identity predicate $\equiv$.

Let us suppose that the absolutist and relativist accept the terms of the dialectical challenge proposed by Williamson (2003, pp. 434–5). The absolutist stakes his position on a particular perspective $\boldsymbol{p}$. He does his best to remove any contextual restrictions and to expand the universe as much as is possible and then claims that the domain $\boldsymbol{mm}$ of this perspective is absolute. The relativist then attempts to induce a shift of perspective—call the resulting perspective $\boldsymbol{p'}$, say—and argues that the domain of this perspective is more inclusive than the first one. If the new domain contains something not among $\boldsymbol{mm}$ the absolutist agrees to concede defeat; if not, the relativist loses.

To reason in plural languages we need to deploy a plural logic. The logic we shall use is FPFO, whose axioms and rules are displayed in table 4.1. In order to allow for a smooth integration with the modal operators we shall add later on, the first-order fragment of this logic falls slightly short of classical. FPFO is equipped with the axioms and rules characteristic of a negative free logic. We use a defined singular existence predicate Ev as an 'existential guard' to restrict Universal Specification and to liberalise Universal Generalisation to obtain their free analogues, FUS and FUG. We also restrict the classical reflexivity principle to obtain Refl_E. We define Ev as $v = v$, which may be easily seen to be provably equivalent to $\exists u(u = v)$ (for u distinct from v). The axioms governing plural quantification and identity are the analogues of these principles, to which we add Plural Comprehension. The plural existence predicate Evv is defined as $vv \equiv vv$, which is provably equivalent in this system to $\exists uu(uu \equiv vv)$ (for uu distinct from vv).[15]

[15]For overviews of plural logic see Burgess and Rosen (1997), Rayo (2002) and Linnebo (2013)

Table 4.1: The axioms and rules of FPFO

Taut	$\top$, for any propositional tautology $\top$.
MP	$\dfrac{\phi_1 \quad \phi_1 \to \phi_2}{\phi_2}$
FUS	$\forall v\phi \to (Eu \to \phi[u/v])$
FUG	$\dfrac{Eu \to (\phi_1 \to \phi_2[u/v])}{\phi_1 \to \forall v\phi_2}$
Refl$_E$	$\Phi(v) \to v = v$
Sub	$u = v \to (\phi \to \phi(u/v))$
FUS^{pl}	$\forall vv\phi \to (Euu \to \phi[uu/vv])$
FUG^{pl}	$\dfrac{Euu \to (\phi_1 \to \phi_2[uu/vv])}{\phi_1 \to \forall vv\phi_2}$
Refl$^{pl}_E$	$\Phi(vv) \to vv \equiv vv$
Sub^{pl}	$uu \equiv vv \to (\phi \to \phi(uu/vv))$
PC	$\exists vv\forall v(v \prec vv \leftrightarrow \phi)$

The axioms and rules are subject to the following side conditions: in FUG (FUG^{pl}) u (uu) does not occur free in the conclusion; in PC, vv does not occur free in ϕ and in Refl$_E$ (Refl$^{pl}_E$) v (vv) *does* occur free in the atomic sentence $\Phi(v)$.

Table 4.2: Semantics for perspective p.

Let $\text{Den}_\sigma(vv)$ abbreviate 'the zero or more objects assigned to vv by σ'; $\text{Den}_\sigma(\langle\lambda_q v.\phi\rangle)$ abbreviate 'the zero or more members of M_q such that ϕ is true relative to q under σ'.

$u = v$ is true relative to p under σ iff $\sigma(u)$ is identical to $\sigma(v)$, which is a member of M_p

$\tau\tau_1 \equiv \tau\tau_2$ is true relative to p under σ iff $\text{Den}_\sigma(\tau\tau_1)$ are identical to $\text{Den}_\sigma(\tau\tau_2)$, which are members of M_p

$u \prec \tau\tau$ is true relative to p under σ iff $\sigma(u)$ is one of $\text{Den}_\sigma(\tau\tau)$, which are members of M_p

$\neg\phi$ is true relative to p under σ iff ϕ is not true relative to p under σ.

$\phi_1 \wedge \phi_2$ is true relative to p under σ iff ϕ_1 is true and ϕ_2 is true relative to p under σ.

$\forall v\phi$ is true relative to p under σ iff, for every a in D_p, ϕ is true relative to p under $\sigma[v/a]$.

$\forall vv\phi$ is true relative to p under σ iff, for all zero or more aa in D_p, ϕ is true relative to p under $\sigma[vv/aa]$.

The reason for the retreat to free axioms and rules for the singular quantifier $\forall v$ is that we sometimes need to allow singular variables to be assigned to objects outside its domain. Recall that we take it to be partially constitutive of what a perspective is that logical vocabulary receives its intended interpretation over the domain D_p of each perspective p; so that the usual satisfaction conditions over D_p displayed in table 4.2 hold relative to p.[16] What calls for a free logic is our continuing to take the satisfaction clauses to hold for assignments σ *not* over the domain D_p. Consequently, to infer $\phi[u/v]$ on the basis of $\forall v\phi$ we need to ensure that the variable u denotes a member of the domain of the quantifier $\forall v$. Provided that the extension of the existential guard Ev comprises this domain, this need is met by the addition of Ev as a further antecedent condition to the classical version of universal specification:

(US) $$\forall v\phi \rightarrow \phi[u/v]$$

Without this restriction, instances of this axiom may fail to be true. If every member of the domain D_p satisfies ϕ but some object a outside D_p fails to, the antecedent is true and the consequent false when u is assigned to a. Similarly to infer $\phi[uu/vv]$ on the basis of $\forall vv\phi$ we need to ensure that the variable uu does not denote any thing outside the domain.

The use of negative rules, embodied in the restriction in the identity axioms Refl_E and Refl_E^{pl} and the definition of the existence predicates Ev and Evv, reflects an actualist attitude to atomic predication. Each satisfier of an atomic predicate relative to a perspective is a member of the domain. Consequently, no modification needs to be made to the satisfaction clause for atomic sentences when we allow for variables to be assigned to objects outside the domain.

The rationale here is that we take the extension of each atomic predicate to comprise every satisfier of that predicate (as interpreted relative to that perspective). The extension is not a restriction of the predicate's full extension to a given domain. The reason extensions of unary predicates are included within the domain reflects domains

and for free logic see Bencivenga (1986) and Lambert (2001).

[16]As in chapter 3, the absolutist can encode assignments using plural term denotations.

Table 4.3: The lexicon of $\mathcal{L}_1^{pl}$

Category e	singular variables: $x, y, z, \ldots$
Category ee	plural variables: $xx, yy, zz, \ldots$
Category $t/(e, e)$	the singular identity predicate: $=$
Category $t/(ee, ee)$	the plural identity predicate: $\equiv$
Category $t/(e, ee)$	the member-'plurality' predicate: $\prec_2$
Category $t/(e, e)$	the element-set predicate: $\in_2$
Category t/e	set predicates: ß$_1$ and ß$_2$
Category t/t and $t/(t, t)$	connectives: $\neg$ and $\wedge$
Category t/t	singular quantifiers: $\forall_2 x$, $\forall_2 y$, $\forall_2 z$,....
Category t/t	plural quantifiers: $\forall_2 xx$, $\forall_2 yy$, $\forall_2 zz$,....
Category ee	plural term: $\boldsymbol{mm}$

being determined from extensions rather than extensions restricted to domains. At least this is the reason for the expansionist. The domain of a perspective is always unrestricted: it is the universe of the lexicon as interpreted relative to the interpretation supplied by the perspective. But the universe is determined, *inter alia* by the extension of predicates. Any object in the extension of a predicate is *ipso facto* a member of the universe and consequently, on the expansionist account, a member of the domain. The contextual restrictionist might offer a similar account. Domains of perspectives may be background domains (in the sense of Glanzberg (2006, p. 49)), encompassing every object in the extension of any predicate relative to the context.[17]

Having modified the classical rules in this way, it is readily verified that, relative to each perspective, each axiom is true under *every* assignment and each rule preserves truth-under-every-assignment. All of this will no doubt appear needlessly complicated and baroque to the absolutist. For, so long as the relativist's shift of perspective does not backfire by actually reducing the extent of the domain, he contends that the domain of $\boldsymbol{p}'$ like that of $\boldsymbol{p}$ is ultimate. Consequently, there are no assignments to objects *outside* this domain. Nonetheless, so long as he accepts the classical principles and their extension to plural quantifiers, along with Plural Comprehension, there is nothing here that he disagrees with. For anything derivable in FPFO is derivable in the stronger system. The absolutist's commitment to this plural logic also commits him to FPFO.

With the logic in hand, let us turn to the argument itself. The argument takes place

[17]See Burge (1974) for a rather different defence of negative free logic.

in the language $\mathcal{L}_1^{pl}$, whose lexicon is displayed in table 4.3. As mentioned above, the argument takes place after the attempted shift in perspective. To remind ourselves of this we attach subscripts to the predicates of the plural language of set theory with urelements, $\mathcal{L}_1^{pl}$: writing ß, $\in$ and $\forall$ as $ß_2$, $\in_2$ and $\forall_2$, and so on, when these expressions receive their intended interpretation over the domain of the second perspective $\boldsymbol{p'}$. (For the sake of concreteness the intended interpretation of ß is pure or impure set;[18] although the arguments from paradox could equally be run on pure sets.) We suppose moreover that we have some metalinguistic vocabulary available, allowing us to describe the initial interpretation from the new perspective. Informally we shall call those objects that satisfy ß relative to the initial perspective, initial sets or sets_1 and those objects that satisfy ß relative to the new perspective new sets or sets_2. But we also suppose that the object language is equipped with metalinguistic expressions: a plural term $\boldsymbol{mm}$ and a predicate $ß_1$, which attempt to encode the initial domain and extension of ß from the new perspective. Of course, in the event that the relativist's attempted shift backfires and makes the domain smaller, we may be unable to refer to everything in the initial universe relative to the new perspective. Still for the purpose of the first argument, it will suffice to have $\boldsymbol{mm}$ denote the intersection of the new and initial domains. Likewise, the intended extension of $ß_1$ comprises the initial sets in the new domain.

Provided that the absolutist holds his nerve he will contend from the new perspective that no expansion of the domain has been achieved: everything in the new domain is in the initial one, and thus in their intersection $\boldsymbol{mm}$. Thus he upholds the following.

Comprehensive Initial Domain. $\forall_2 x(x \prec_2 \boldsymbol{mm})$

Unlike Comprehensive Domain this is no longer a trivial truth.

The relativist's argument against this claim has two premisses. The first is a restricted version of Collapse, expressing a maximally liberal attitude to the formation of new sets from initial ones. Define a plural predicate B_1vv (B_2vv) which applies to

[18]A set is said to be *pure* if there are no urelements (non-sets) in its transitive closure: that is, if no urelement is an element of the set or an element of an element, or an element of an element of an element, and so on; otherwise it is said to be *impure*.

some things just in case each of them in the new domain is a set$_1$ (set$_2$):

$$B_1xx =_{\mathrm{df}} \forall_2 x(x \prec_2 xx \rightarrow \text{ß}_1 x)$$

$$B_2xx =_{\mathrm{df}} \forall_2 x(x \prec_2 xx \rightarrow \text{ß}_2 x)$$

And let the singular and plural quantifiers $\forall_1 v$ and $\forall_1 vv$ be restricted to the initial domain $\boldsymbol{mm}$ as follows.

$$\forall_1 v\phi =_{\mathrm{df}} \forall_2 v(v \prec_2 \boldsymbol{mm} \rightarrow \phi)$$

$$\forall_1 vv\phi =_{\mathrm{df}} \forall_2 vv\big(\forall_2 v(v \prec_2 vv \rightarrow v \prec_2 \boldsymbol{mm}) \rightarrow \phi\big)$$

The first premiss may then be stated as follows.

Liberalness (*Metalinguistic*). For any initial sets (in the initial domain) some new set has them as elements (strictly, elements$_2$)

$\forall_1 xx\big(B_1xx \rightarrow \exists_2 y(\text{ß}_2 y \wedge \forall_2 z(z \in_2 y \leftrightarrow z \prec_2 xx))\big)$

The second premiss states that any shift in the extension of ß is 'stable': the only potential change to the extension of ß is the addition of new objects, that is, objects which do not fall in the initial domain.

Atomic Stability (*Metalinguistic*) For any member of the initial domain: (S^+) if it is an initial set, it is a new set and (S^-) if it is a new set, it is an initial set.

(S^+) $\forall_1 x(\text{ß}_1 x \rightarrow \text{ß}_2 x)$

(S^-) $\forall_1 x(\text{ß}_2 x \rightarrow \text{ß}_1 x)$

Together with the absolutist's thesis, Comprehensive Initial Domain, it is routine to derive a contradiction from the two premisses in FPFO. Given the absolutist's claim, the restriction to $\boldsymbol{mm}$ in $\forall_1 v$ and $\forall_1 vv$ is vacuous. Consequently, each restricted quantifier $\forall_1$ in Liberalness and Atomic Stability may be replaced with the unrestricted $\forall_2$. The new version of Atomic Stability ensures the coextensiveness of $\text{ß}_2 v$ and $\text{ß}_1 v$, and thereby also ensures the coextensiveness of the defined plural predicates $B_2 vv$ and $B_1 vv$. Consequently every subscript 1 in Liberalness may be replaced with 2, and we arrive at a mere notational variant of Collapse, restricted to sets.

Collapse $\forall_2 xx\big(B_2 xx \to \exists_2 y(ß_2 y \wedge \forall_2 z(z \in_2 y \leftrightarrow z \prec_2 xx))\big)$

This leads to a contradiction in FPFO by the Russell argument, using the instance of Plural Comprehension for the condition $ß_2 x \wedge \neg x \in_2 x$. Consequently, the negation of the absolutist thesis may be derived from Liberalness and Atomic Stability in FPFO

Non-Comprehensive Initial Domain. $\neg\forall_2 x(x \prec_2 \boldsymbol{mm})$

As such, if the premisses can be sustained, the absolutist loses the dialectical challenge.[19]

We shall return to defend these premisses in sections 3 and 4 below. But even granted these premisses, the argument remains a less than satisfying response to absolutism. For though the absolutist loses the dialectical challenge, this falls short of a refutation of absolutism, as we saw in chapter 3 (section 1.1). It is still open to him to claim, albeit with rather reduced dialectical credibility, that another domain—even the domain of the new perspective $\boldsymbol{p}'$—is absolute. The relativist should accordingly seek to establish a stronger conclusion. As we saw in the previous chapter, the schematic and modal resources introduced there provide the means for the relativist to frame such a conclusion. We shall now consider how the relativist can deploy these resources to frame the premisses of the argument from paradox, and thereby arrive at the stronger conclusion.

2.2 The schematic argument

Recall that the schematic characterisation of relativism formulates this theory with infinitely many axioms. Each axiom expresses that the domain of a given perspective—encoded by a plural term of the form $mm_\varsigma =_{\text{df}} \langle\lambda_\varsigma y.\exists x(x = y)\rangle$—fails to be comprehensive from a suitable meta-perspective. The axioms and the perspectives relative to which they are to be interpreted are recorded in the following inter-perspective schema.

Relativism $\varsigma'\colon \neg\forall x(x \prec mm_\varsigma)$

[19]The metalinguistic argument shares a number of similarities with a less explicitly metalinguistic argument advanced by Fine. See (2006, pp. 21–2).

Table 4.4: The lexicon of $\mathcal{L}_2^{pl}$

Category e	singular variables: $x, y, z, \ldots$
Category ee	plural variables: $xx, yy, zz, \ldots$
Category $t/(e,e)$	the singular identity predicate: $=$
Category $t/(ee,ee)$	the plural identity predicate: $\equiv$
Category $t/(e,ee)$	the member-'plurality' predicate: $\prec$
Category $t/(e,e)$	the element-set predicate: $\in$
Category t/e	the set predicate: $ß$
Category t/t and $t/(t,t)$	connectives: $\neg$ and $\wedge$
Category t/t	singular quantifiers: $\forall x$, $\forall y$, $\forall z$,. . . .
Category t/t	plural quantifiers: $\forall xx$, $\forall yy$, $\forall zz$,. . . .
Category ee/t	plural-term-forming operators: λ_p, $\lambda_{p'}$, . . .

The second version of the argument from paradox has this formulation of relativism as its target conclusion. Accordingly the argument is formulated in a language $\mathcal{L}_2^{pl}$ that enriches $\mathcal{L}_{ß,\in}$, with λ-terms of the form $\langle \lambda_\varsigma v.\phi \rangle$. (See table 4.4.) In the metalanguage we use perspective terms: $p, q \ldots$ and perspective variables: $r, s, \ldots$ to label sentences, as in chapter 3 (section 2.2). We no longer require the additional metalinguistic vocabulary deployed in the previous argument, and since shifts in perspective are now indicated with labels we stop indexing the object language predicates: $ß$, $\in$, and so on. The key premisses still recognisably embody a liberal attitude to set formation and a stability assumption, but their schematic formulations differ somewhat from the metalinguistic formulation.

The appeal to Liberalness in the informal argument is more naturally encapsulated as a rule of inference rather than a metalinguistic thesis. The relativist starts off by assuming 'These sets comprise the non-self-members sets' relative to one perspective, p say; then—having shifted to another perspective, p'—concludes 'Some set has those sets as elements' (with 'these sets' and 'those sets' referring to the same sets). We shall record such *inter*-perspective inferences using labelled sequents. The inference may be formalised as:

$$p\colon Err, p\colon Brr \Rightarrow p'\colon \exists y(ßy \wedge \forall z(z \in y \leftrightarrow z \prec rr)$$

where $rr =_{\text{df}} \langle \lambda_p x.x \notin x \rangle$. The arrow $\Rightarrow$ occurs not within a labelled sentence but between labelled sentences.[20]

[20]Strictly, the arrow occurs between two sets of labelled sentences.

More generally, we shall formulate such inter-perspective inferences using labelled sequents.[21] To start with, consider labelled sequents of the form $\Gamma \Rightarrow \varsigma\colon \phi$, where Γ is a set of labelled sentences, and $\varsigma\colon \phi$ a labelled sentence. We shall call each sentence that occurs (prefixed with a label) in Γ an *antecedent sentence*, and the proposition expressed by an antecedent sentence, relative to the labelled perspective (and an assignment σ) an *antecedent proposition* (under σ). The intended reading of such a sequent is as a pattern of inference. The sequent records (correctly or incorrectly) that what is expressed by ϕ relative to the perspective labelled by ς follows from the antecedent propositions.

The metalinguistic formulations of Liberalness and Atomic Stability may each be replaced with inter-perspective schemas. Write Bvv to abbreviate $\forall x(x \prec vv \rightarrow ßx)$, as before. Liberalness may be formulated simply by replacing the perspective constant p in the above sequent with a schematic letter and rr with a plural variable.

Liberalness (*Schema*) $\varsigma\colon E\tau\tau, \varsigma\colon B\tau\tau \Rightarrow \varsigma'\colon \exists y(ßy \wedge \forall z(z \in y \leftrightarrow z \prec \tau\tau))$

subject to the usual side condition that the schematic letter be replaced with a perspective term. Just as inter-perspective sentence schemas allow the relativist to indicate commitments that she is unable to state from her current perspective, inter-perspective sequent schemas allow the relativist to indicate her inferential commitments relative to perspectives she is presently unable to quantify over. As in chapter 3, such schemas should be taken to be open-ended. Their instances are not restricted to perspective terms that denote perspectives that the relativist is currently able to refer to or quantify over in her current language. (See section 2.1.)

The Atomic Stability principle deployed in the previous argument may be generalised in a similar way. We also generalise it to atomic sentences other than those of the form $ßv$. Given any atomic sentence Φ, with free variables $v_1, \ldots, vv_n$, we write E_Φ to abbreviate $Ev_1 \wedge \ldots \wedge Evv_n$, with Ev and Evv defined as above. Provided logical expressions receive their intended interpretations at a perspective, E_Φ states that the parameters in Φ fall within the domain of that perspective: E_Φ is true relative to a

[21]Labelled sequents have recently been the focus of proof theoretic investigation, especially in the contexts of modal systems, for their pleasing proof theoretic properties. See, for instance, Negri (2005). See Viganó (2000) for an overview.

perspective just in case each singular variable v in Φ denotes a member of the domain and each plural variable vv in Φ denotes some zero or more members of the domain (and nothing outside the domain). The stability of Φ between object and meta-perspectives may then be characterised with schematic analogues of the previous principles.

Atomic Stability (*Schemas*)

(S$^+$) $\varsigma : E_\Phi, \varsigma : \Phi \Rightarrow \varsigma' : \Phi$

(S$^-$) $\varsigma : E_\Phi, \varsigma' : \Phi \Rightarrow \varsigma : \Phi$

Whenever Φ's free variables only denote members of the object domain, instances of these two schemas ensure that its truth-value is invariant across the object and meta-perspective: an instance of the first ensures that truth transmits upwards from object to meta-perspective; an instance of the second ensures that it transmits downwards from meta- to object perspective.

Having schematically generalised the premisses and conclusion of the argument from paradox, we also need to generalise our means of reasoning to apply to such labelled formulas and sequents. A labelled version of a Gentzen-style sequent calculus provides the requisite formal tool.[22]

As is usual in this setting, we generalise labelled sequents to permit multiple labelled sentences after the sequent arrow. In full generality, a labelled sequent is of the form $\Gamma \Rightarrow \Theta$ where each of Γ and Θ are sets of labelled sentences. The sentences labelled in Γ are called *antecedent sentences*, as before, and the sentences labelled in Θ, *consequent sentences*. The propositions expressed by such sentence relative to their labelled perspective are called antecedent and consequent propositions.

(We adopt the usual notational conventions for such sequents: set brackets and union symbols are dropped $\Gamma \cup \{\varsigma : \phi\}$ and $\Gamma \cup \Gamma'$ are abbreviated as $\Gamma, \varsigma : \phi$ and Γ, Γ', and so on. Moreover, we drop set terms for the empty set: $\Gamma \Rightarrow \varnothing$ is abbreviated as $\Gamma \Rightarrow$, and so on.)

[22]The sequent calculus and the accompanying semantics presented below adapts Viganò's (2000) presentation of a singular modal calculus to the present plural non-modal setting.

In addition to labelled analogues of the usual structural rules, the sequent version of the argument from paradox will draw on two sorts of logical rules. (See table 4.5 on page 122 below.) Rules of the first sort embody a labelled sequent formulation of negative free plural first-order logic for the language in question. These make two modifications to the standard Gentzen operational rules for connectives, singular and plural quantifiers and λ-terms. First, the sentences are labelled. Second, existential guards are added where necessary, as before. These rules are all intra-perspective insofar as sentences other than 'side formulas' labelled in Γ and Θ are all labelled by the same perspective term. The second sort of rules are inter-perspective. They govern the interaction of the existence predicates and the member-'plurality' relation $\prec$ across different perspectives. These rules reflect the fact that plural variables are taken to be rigid designators: under an assignment, vv denotes the same things relative to each perspective.

We now have everything in place for the second version of the argument from paradox to be made. Much as before, in conjunction with Atomic Stability, the assumption that the meta-domain is no wider than the object domain—now articulated with what is expressed by $\forall x(x \prec mm_\varsigma)$ at the meta-perspective labelled by ς'—allows us to show that the truth-value of Bxx is invariant across object and meta-perspective. Indeed such invariance may be established for any sentence from this assumption (see the appendix for a proof).

Invariance

(I$^+$) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi \Rightarrow \varsigma'\colon \phi$

(I$^-$) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \phi \Rightarrow \varsigma\colon \phi$

The relativist then argues as follows to establish the schematic formulation of relativism on the basis of the Invariance of Exx and Bxx and Liberalness.[23]

[23] We avail ourselves of the usual free derived rules for connectives and quantifiers other than those in the official lexicon.

Table 4.5: A sequent formulation of plural logic

Structural rules

$$\textsc{rep}\ \frac{}{\varsigma\colon\phi, \Gamma \Rightarrow \Theta, \varsigma\colon\phi} \qquad\qquad \frac{\varsigma\colon\phi, \Gamma \Rightarrow \Theta \qquad \Gamma \Rightarrow \Theta, \varsigma\colon\phi}{\Gamma \Rightarrow \Theta}\ \textsc{cut}$$

Intra-perspective rules

$$\textsc{l}\neg\ \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\phi}{\varsigma\colon\neg\phi, \Gamma \Rightarrow \Theta} \qquad\qquad \frac{\varsigma\colon\phi, \Gamma \Rightarrow \Theta}{\Gamma \Rightarrow \Theta, \varsigma\colon\neg\phi}\ \textsc{r}\neg$$

$$\textsc{l}\wedge\ \frac{\varsigma\colon\phi_1, \varsigma\colon\phi_2, \Gamma \Rightarrow \Theta}{\varsigma\colon\phi_1 \wedge \phi_2, \Gamma \Rightarrow \Theta} \qquad\qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\phi_1 \qquad \Gamma \Rightarrow \Theta, \varsigma\colon\phi_2}{\Gamma \Rightarrow \Theta, \varsigma\colon\phi_1 \wedge \phi_2}\ \textsc{r}\wedge$$

$$\textsc{l} =\ \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\phi}{\varsigma\colon\tau_1 = \tau_2, \Gamma \Rightarrow \Theta, \varsigma\colon\phi(\tau_2/\tau_1)} \qquad\qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\Phi(\tau)}{\Gamma \Rightarrow \Theta, \varsigma\colon\tau = \tau}\ \textsc{r} =$$

$$\textsc{l} =\ \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\phi}{\varsigma\colon\tau\tau_1 \equiv \tau\tau_2, \Gamma \Rightarrow \Theta, \varsigma\colon\phi(\tau\tau_2/\tau\tau_1)} \qquad\qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\Phi(\tau\tau)}{\Gamma \Rightarrow \Theta, \varsigma\colon\tau\tau \equiv \tau\tau}\ \textsc{r} =$$

$$\textsc{l}\lambda\ \frac{\varsigma\colon E\tau, \varsigma\colon\phi[\tau/v], \Gamma \Rightarrow \Theta}{\varsigma\colon\tau \prec \langle\lambda_\varsigma v.\phi\rangle, \Gamma \Rightarrow \Theta} \qquad\qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon\phi[\tau/v] \qquad \Gamma \Rightarrow \Theta, \varsigma\colon E\tau}{\Gamma \Rightarrow \Theta, \varsigma\colon\tau \prec \langle\lambda_\varsigma v.\phi\rangle}\ \textsc{r}\lambda$$

$$\textsc{l}\forall v\ \frac{\varsigma\colon\phi[\tau/v], \Gamma \Rightarrow \Theta \qquad \Gamma \Rightarrow \Theta, E\tau}{\varsigma\colon\forall v\phi, \Gamma \Rightarrow \Theta} \qquad\qquad \frac{\varsigma\colon Eu, \Gamma \Rightarrow \Theta, \varsigma\colon\phi[u/v]}{\Gamma \Rightarrow \Theta, \varsigma\colon\forall v\phi}\ \textsc{r}\forall v$$

$$\textsc{l}\forall vv\ \frac{\varsigma\colon\phi[\tau\tau/vv], \Gamma \Rightarrow \Theta \qquad \Gamma \Rightarrow \Theta, E\tau\tau}{\varsigma\colon\forall vv\phi, \Gamma \Rightarrow \Theta} \qquad\qquad \frac{\varsigma\colon Euu, \Gamma \Rightarrow \Theta, \varsigma\colon\phi[uu/vv]}{\Gamma \Rightarrow \Theta, \varsigma\colon\forall vv\phi}\ \textsc{r}\forall vv$$

$$\frac{}{\Rightarrow \varsigma\colon E\langle\lambda_\varsigma v.\phi\rangle}\ \mathrm{PC}$$

Inter-perspective rules

$$\textsc{l}Evv\ \frac{}{\varsigma\colon E\tau\tau, \varrho\colon v \prec \tau\tau \Rightarrow \varsigma\colon Ev} \qquad\qquad \frac{v\colon u \prec \tau\tau, \Gamma \Rightarrow \Theta, \varsigma\colon Eu}{\Gamma \Rightarrow \Theta, \varsigma\colon E\tau\tau}\ \textsc{r}Evv$$

Side constraints: in R$\forall v$ (R$\forall v$) the variable u (uu), does not occur free in the end sequent; in R $=$, the term τ does occur free in $\Phi(\tau)$; in the right rule for REvv neither the object variable u nor the perspective variable v occurs free in the end sequent.

$$
\begin{array}{c}
(\mathrm{I}^{-}) + \text{Liberalness} \\
\vdots \\
\varsigma'\colon \forall x(x \prec mm_{\varsigma}), \varsigma'\colon Exx, \varsigma'\colon Bxx \;\Rightarrow\; \varsigma'\colon \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec xx)) \\
\hline
\varsigma'\colon \forall x(x \prec mm_{\varsigma}), \varsigma'\colon Exx \;\Rightarrow\; \varsigma'\colon Bxx \rightarrow \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec xx)) \\
\hline
\varsigma'\colon \forall x(x \prec mm_{\varsigma}) \;\Rightarrow\; \varsigma'\colon \forall xx(Bxx \rightarrow \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec xx))) \\
\hline
\vdots \\
\text{Russell argument} \\
\vdots \\
\hline
\varsigma'\colon \forall x(x \prec mm_{\varsigma}) \;\Rightarrow \\
\hline
\Rightarrow\; \varsigma'\colon \neg\forall x(x \prec mm_{\varsigma})
\end{array}
$$

The first lacuna is to be filled with the derivation of the instance of (I^{-}) for Exx and Bxx, followed by two applications of cut to replace the labelled antecedent sentences $\varsigma\colon Exx$ and $\varsigma\colon Bxx$ in the sequent formulation of Liberalness with $\varsigma'\colon Exx$ and $\varsigma'\colon Bxx$. Having reached what is simply a labelled version of the Collapse principle to the right of the sequent arrow, the second lacuna is filled by carrying out the usual refutation of this principle in free plural first-order logic as follows. Write rr to abbreviate $\langle \lambda_{\varsigma'} z.\text{ß}z \wedge z \notin z\rangle$. Then we reason as follows.

$$
(\Pi_1)\qquad
\dfrac{
\dfrac{
\dfrac{
\dfrac{\varsigma'\colon y \notin y \Rightarrow \varsigma'\colon y \notin y \qquad \varsigma'\colon \text{ß}y \Rightarrow \varsigma'\colon \text{ß}y}{\varsigma'\colon \text{ß}y, \varsigma'\colon y \notin y \Rightarrow \varsigma'\colon \text{ß}y \wedge y \notin y}
\qquad \varsigma'\colon \text{ß}y \Rightarrow \varsigma'\colon Ey}
{\varsigma'\colon \text{ß}y, \varsigma'\colon y \notin y \Rightarrow \varsigma'\colon y \prec rr}}
{\varsigma'\colon \text{ß}y \Rightarrow \varsigma'\colon y \in y, \varsigma'\colon y \prec rr}
\qquad
\dfrac{
\dfrac{
\dfrac{\varsigma'\colon y \notin y \Rightarrow \varsigma'\colon y \notin y}{\varsigma'\colon \text{ß}y \wedge y \notin y \Rightarrow \varsigma'\colon y \notin y}}
{\varsigma'\colon y \prec rr \Rightarrow \varsigma'\colon y \notin y}}
{\varsigma'\colon y \prec rr, \varsigma'\colon y \in y \Rightarrow}
}{\varsigma'\colon \text{ß}y, \varsigma'\colon y \in y \leftrightarrow y \prec rr \Rightarrow}
$$

Second, it straightforward to demonstrate Brr $(=_{\mathrm{df}} \forall x(x \prec rr \rightarrow \text{ß}x))$, relative to ς'

$$
(\Pi_2)\qquad
\begin{array}{c}
\varsigma'\colon \text{ß}x \Rightarrow \varsigma'\colon \text{ß}x \\
\hline
\varsigma'\colon \text{ß}x \wedge x \notin x \Rightarrow \varsigma'\colon \text{ß}x \\
\hline
\varsigma'\colon x \prec rr \Rightarrow \varsigma'\colon \text{ß}x \\
\hline
\Rightarrow \varsigma'\colon x \prec rr \rightarrow \text{ß}x \\
\hline
\Rightarrow \varsigma'\colon Brr
\end{array}
$$

Third, we reason as follows.

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\Pi_1\\ \vdots\\ \varsigma'\!: \text{ß}y, \varsigma'\!: y \in y \leftrightarrow y \prec rr \Rightarrow\end{array} \qquad \varsigma'\!: Ey \Rightarrow \varsigma'\!: Ey}{\varsigma'\!: Ey, \varsigma'\!: \text{ß}y, \varsigma'\!: \forall z(z \in y \leftrightarrow z \prec rr) \Rightarrow}}{\varsigma'\!: Ey, \varsigma'\!: \text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec rr) \Rightarrow}}{\varsigma'\!: \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec rr)) \Rightarrow} \qquad \begin{array}{c}\Pi_2\\ \vdots\\ \Rightarrow \varsigma'\!: Brr\end{array}}{\varsigma'\!: Brr \rightarrow \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec rr)) \Rightarrow} \qquad \Rightarrow \varsigma'\!: Err}{\varsigma'\!: \forall xx(Bxx \rightarrow \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec xx))) \Rightarrow}
$$

Finally, we apply CUT. With the gaps filled in, the result is a proof whose only schematic letter is ς. Its conclusion is the schematic version of Relativism and its premisses are schematic versions of our two key premisses.

What is the absolutist to make of such a proof-schema? As mentioned before, we shall leave the defence of the key sequents, the Liberalness and Atomic Stability sequents until sections 3 and 4. What I want to argue here is first that the argument is a threat to absolutism and second that the only viable response for the absolutist—or, at least for those absolutists who do not object to the semantics in table 4.2—is to reject one or other of the key sequents.

The schematic proof indicates how the relativist will argue against the absolutist relative to any perspective. Given any perspective $\boldsymbol{p}$—including any whose domain is claimed to be absolute—the sentence labelled in the conclusion of the end sequent when the perspective term $\boldsymbol{p}$ replaces the schematic letter is $\neg\forall x(x \prec mm_{\boldsymbol{p}})$. If true at the perspective $\boldsymbol{p}'$, this ensures that the domain of $\boldsymbol{p}$ fails to be absolute. In the normal case an argument ensures the truth of its conclusion if each premiss is true and each rule of inference truth-preserving. For a sequent derivation, since each step moves between patterns of inference rather than object language sentences a slightly different standard is required. First, we need to generalise the notion of assignment. An assignment σ over the perspectives p_0, p_1, ... assigns (i) each singular object variable v to a member of the domain D_r of some perspective r among p_0, p_1, ... (ii) each plural object variable vv to some zero or more members of the domain D_r (and nothing else) for some perspective r among p_0, p_1, ...; and (iii) each perspective variable v to one of the perspectives p_0, p_1, ... We then say that a sequent $\Gamma \Rightarrow \Theta$ is *sound* over the perspectives p_0, p_1, ... if there is no assignment σ over p_0, p_1, ... such that each antecedent sentence is true relative to the perspective denoted by its label under σ and each consequent sentence is false relative to the perspective denoted by its label under σ. But granted that the semantic clauses displayed in table 4.2 hold for $\boldsymbol{p}$ replacing p and for $\boldsymbol{p}'$ replacing p, it is readily verified that the logical sequents and sequent rules displayed in table 4.5 are sound or preserve soundness over $\boldsymbol{p}$ and $\boldsymbol{p}'$. The non-absoluteness of the domain of $\boldsymbol{p}$ consequently comes down to whether or not instances of the Liberalness and Atomic Stability sequents are sound over these perspectives. If so, the end sequent in the proof is sound over this domain. And since it lacks an antecedent sentence, this ensures that its conclusion $\neg\forall x(x \prec mm_{\boldsymbol{p}})$ is true relative to $\boldsymbol{p}'$.

2.3 The modal argument

The schematic formulation of the argument from paradox improves on the metalinguistic formulation in terms of generality. Where the metalinguistic argument failed to rule out the absoluteness of domains wider than $\boldsymbol{mm}$, any domain's absoluteness is denied by some instance of the schematic formulation of relativism. Nonetheless the relativist may find yet further generality desirable. As we observed in the previous chapter, the schematic formulation of relativism is more like an undertaking to offer partial formulations of the view in certain circumstances than an actual *statement* of relativism. The schema template, by itself, does not state anything, for it contains uninterpreted schematic letters, or placeholders. Rather it indicates the form of many sentences, each of which states part of the view. The argument in the previous section likewise schematically indicates the form of the premisses the relativist can state and the inferences she can deploy in order to reach a partial statement of the view, denying the comprehensiveness of a single domain. But she is unable to fill in the schematic blanks, and actually state the premisses and conclusions of the relevant instance, until she has acquired the expressive power to do so. She cannot deny the comprehensiveness, and thus absoluteness, of a given domain until she has reached a suitable meta-perspective.

Pluralist relativists can offer a further improvement in the statement of the argument from paradox. Recall that relativists of this stripe allow that while quantifiers themselves cannot achieve absolute generality, such generality can be achieved by using

Table 4.6: The lexicon of $\mathcal{L}_3^{pl}$

Category e	singular variables: $x, y, z, \ldots$
Category ee	plural variables: $xx, yy, zz, \ldots$
Category $t/(e, e)$	the singular identity predicate: $=$
Category $t/(ee, ee)$	the plural identity predicate: $\equiv$
Category $t/(e, ee)$	the member-'plurality' predicate: $\prec$
Category $t/(e, e)$	the element-set predicate: $\in$
Category t/e	the set predicate: *ß*
Category t/t and $t/(t, t)$	connectives: $\neg$ and $\wedge$
Category t/t	singular quantifiers: $\forall x$, $\forall y$, $\forall z$,. . . .
Category t/t	plural quantifiers: $\forall xx$, $\forall yy$, $\forall zz$,. . . .
Category t/t	an inter-perspective modal operator: $\Box$

a suitable inter-perspective modal operator, $\Box$. The third regimentation of the argument from paradox takes place in a language that adds $\Box$ to plural language of set theory with urelements. (See table 4.6.)

As we saw in the previous chapter such modal resources allow the pluralist to state her view using a single modal sentence.

Relativism (modal statement). $\Box\forall xx\neg\Box\forall x\Diamond(x \prec xx)$

This thesis will be the conclusion of our third formulation of the argument from paradox. And just as modal resources permit a non-schematic statement of the conclusion of this argument, they also permit a non-schematic statement of its premisses. In this way, the pluralist can replace the schematic description of the arguments she could give in various circumstances with a single argument whose premisses may be stated, and whose single conclusion, if true, rules out the absoluteness of any domain.

The third argument draws on broadly the same key premisses as the first two. We revert to regimenting Liberalness and Stability as principles rather than rules of inference. Liberalness is stated as a modal analogue of Collapse. Given a non-modal sentence ϕ let its modal analogue, or *modalisation*, be the modal sentence $\phi^{\Diamond}$ that replaces each quantifier $\forall/\exists$ in ϕ with $\Box\forall/\Diamond\exists$ and prefixes each atomic subsentence in ϕ with $\Diamond$.[24] The modal version of Liberalness is the modalisation of Collapse. (To tidy the notation, we write the prefixed atomic subsentence $\Diamond x \in y$, as $x \in^{\Diamond} y$; similarly for

[24]This is a slight modification of a modal translation investigated by Linnebo in (2010), from whom we also borrow the term 'modalisation'. See chapter 5 n. 25, below.

other such subsentences.)

Liberalness (*Modal*). $\Box\forall xx(B^{\Diamond}xx \rightarrow \Diamond\exists y(\text{ß}^{\Diamond}y \wedge \Box\forall z(z \in^{\Diamond} y \leftrightarrow z \prec^{\Diamond} xx)))$

This is to be sharply distinguished from its inconsistent analogue. Rather than claiming impossibly that any sets in the present domain form a set in that domain, it is the modal generalisation of the weaker claim that any sets in any domain form a set in some other domain.

Atomic Stability may be formulated as a modal generalisation of the claim that, for any assignment to its free variables, $v_1, \ldots vv_n$, if the atomic sentence Φ is true, it remains true relative to any perspective where its parameters exist; and if Φ is false, relative to a perspective at which its parameters exist, it is false relative to every perspective.[25]

Atomic Stability (*Modal*).

(S$^+$) $\Box\forall v_1 \ldots \Box\forall vv_n\big(\Phi \rightarrow \Box(E_\Phi \rightarrow \Phi)\big)$

(S$^-$) $\Box\forall v_1 \ldots \Box\forall vv_n\big(\neg\Phi \wedge E_\Phi \rightarrow \Box\neg\Phi\big)$

As ever, in addition to the premisses, the argument relies on a suitable logic. The language we deploy simply adds $\Box$ to the plural language of set theory, $\mathcal{L}^{pl}_{\text{ß},\in}$. In adding further vocabulary the intended interpretation of the original lexicon remains unchanged. Accordingly, this interpretation continues to sustain FPFO; and indeed the schemas of this logic may be extended to modal sentences. This logic is then extended with principles governing the modal operator, and its interactions with the quantifiers.

Recall that to a first approximation, the intended reading of the modal operator is given by the gloss of $\Box\psi$ as saying that ψ holds relative to every perspective. (Although the pluralist relativist will insist that this gloss is faithful only if the right-hand side is taken to have implicit modal content, with a silent $\Box$ before 'every'.) In omitting mention of an accessibility relation in the truth-condition, we in effect take the underlying accessibility relation to be the universal relation on perspectives. This motivates the addition of the propositional axioms and rules characteristic of S5 to FPFO, with the

[25] Parsons (1983) considers such principles which he calls existence-dependent rigidity principles in the context of developing a modal set theory of the sort that will concern us in the next chapter. We borrow the label 'Stability' from similar principles deployed by Linnebo (n.d.).

Table 4.7: The axioms and rules of $\mathsf{S5}$

K	$\Box(\psi_1 \to \psi_2) \to (\Box\psi_1 \to \Box\psi_2)$
T	$\Box\psi \to \psi$
4	$\Box\psi \to \Box\Box\psi$
B	$\psi \to \Box\Diamond\psi$
Nec	$\dfrac{\psi}{\Box\psi}$

axioms (T), (4) and (B) corresponding to the reflexiveness, transitivity and symmetry of the accessibility relation, as usual. (See table 4.7)[26]

The importance of free formulations of the quantification axioms becomes very clear when we add modal resources to the language. If we add the innocuous-looking axiom Ev, to the $\mathsf{FPFO} + \mathsf{S5}$, then the classical axioms for the first-order quantifiers swiftly follow, and in their wake we may derive the Barcan Formula and its converse.[27]

(BF) $$\forall v \Box\psi \to \Box\forall v\psi$$

(CBF) $$\Box\forall v\psi \to \forall v\Box\psi$$

However, the relativist will not accept these principles. Contrary to her view, the two schemas ensure that domains remain constant from perspective to perspective.

As before, the trouble is with Universal Specification, and now also with the axiom Ev that permits its derivation in the free system. Both principles *are* true relative to each perspective under every assignment *over its domain.* In particular if $\forall v$ expresses unrestricted universal quantification over the domain of the present version of English, whatever we assign to u falls within the range of $\forall v$. But while true under each assignment, instances need not be necessarily true. Consequently, under some assignments, some theorems of classical logic fail to be necessary truths. But once we have a necessitation rule, we can expect a system which proves contingent truths to deliver outright falsehoods. And this is exactly what happens in the case of the derivation of

[26] For overviews of quantified modal logic see, for instance, Hughes and Cresswell (1996), Fitting and Mendelsohn (1998), Braüner and Ghilardi (2007). The axiomatisation of $\mathsf{FPFO}+\mathsf{S5}$ adapts the positive free singular axiomatisation of modal logic offered by Braüner and Ghilardi to the present setting.

[27] See, for instance Hughes and Cresswell (1996, p. 245–7).

the Converse Barcan formula. The schema may be derived as follows.

$\vdash \forall v\phi \to \phi[u/v]$	US
$\vdash \Box(\forall v\phi \to \phi[u/v])$	Nec
$\vdash \Box\forall v\phi \to \Box\phi[u/v]$	K, MP
$\vdash \Box\forall v\phi \to \forall v\Box\phi$	UG

The relativist will object that the transition from the first line to the second fails to preserve truth. For example, when we assign y to the Russell set for some less inclusive perspective, $\{x \in D_p \mid x \notin x\}$, then the instance of universal specification $\forall y\exists x(x = y) \to \exists x(x = y)$ is a trivial truth relative to present perspective but its necessitation $\Box(\forall y\exists x(x = y) \to \exists x(x = y))$ is false. Relative to the smaller perspective p, the antecedent remains true, but no member of the domain D_p is identical to the Russell set for this domain.[28]

While the pluralist relativist rejects the addition of Ev as an axiom, along with that of the Barcan formula and its converse, she does accept some weaker principles. In the first place, she always accepts $\Diamond Ev$, regardless of what v is assigned to. For, she contends, while v may be assigned to objects outside a given perspective's domain, it must be assigned to an object in *some* perspective's domain. Likewise while she rejects Evv she accepts the weaker $\Diamond Evv$. Second, she accepts the following restricted version of the Barcan formula.

(RBF) $$\forall xx\big((\forall y \prec^{\Diamond} xx)\Box\psi \to \Box(\forall y \prec^{\Diamond} xx)\psi\big)$$

In this principle the singular quantifier is restricted, relative to each perspective, to some zero or more things in the domain of the present perspective. But just as the Barcan formula ensures that the domain of the unrestricted quantifier $\forall v$ never increases, the instances of this schema ensure that the domain of the restricted quantifier $(\forall v \prec^{\Diamond} vv)$

[28] Under the metaphysical necessity reading of $\Box$, see Garson (2001) for a defence of modal logic based on free axioms and Williamson (1998a) for a defence of the Barcan formula and its Converse.

never increases. And the relativist maintains that this is the case. For recall that plural variables are taken to be rigid, denoting the same zero or more things, relative to each perspective. Consequently, membership of the things denoted by vv remains constant, and thus never increases. Likewise the relativist accepts the plural analogue of the restricted Barcan formula. Write $vv \prec uu$ to abbreviate $\forall u(u \prec vv \rightarrow u \prec uu)$. (Read: uu is a 'sub-plurality' of vv.) Then the following principle also holds in virtue of the domain of a plural quantifier restricted to some members of the present domain never increasing.

$$\text{(RBF}_p) \qquad \forall yy\big((\forall yy \prec^{\Diamond} xx)\Box\psi \rightarrow \Box(\forall yy \prec^{\Diamond} xx)\psi\big)$$

The third argument will take place against the background of the modal logic $\mathsf{FPFO} + \mathsf{S5} + \Diamond Ev + \Diamond Evv + \text{RBF} + \text{RBF}_p$.

As before there is nothing in the logic that should be objectionable to the absolutist. Provided he concedes the intelligibility of the modal operator, the structure of perspectives that the logic describes is perfectly compatible with his view. In particular, no assumption is made in the logic that domains increase. Of course, the absolutist may find the appeal to modality needlessly baroque. For he maintains that he can achieve with plain quantifiers what the pluralist seeks to achieve by prefixing them with modal operators. Still mere complication is a price worth paying for a language which permits a formulation of the view at issue that is acceptable to both pluralist relativists and absolutists.

But, as ever, while the underlying logic is neutral between absolutism and relativism, the premisses of the argument from paradox permit us to establish relativism. This time rather than showing that sentences are coextensive between an object and meta-perspective, we show that given the absolutist assumption that the domain is absolute, the distinction between a sentence and its modalisation collapses; modalised sentences 'flatten out' relative to an absolute perspective. (See the appendix for a proof.)

Flattening

(F) $\exists xx\Box\forall y(y \prec^{\Diamond} xx), (\text{S}^-), \vdash \phi \leftrightarrow \phi^{\Diamond}$

Applying Flattening to the Liberalness premiss, articulated by *Collapse*$^\Diamond$, permits us to derive the inconsistent principle, *Collapse*. Relativism may then be shown to be derivable from Stability and Liberalness as follows.

$$
\begin{array}{rll}
(\mathrm{S}^-), \mathit{Collapse}^\Diamond, \exists xx \Box \forall y (y \prec^\Diamond xx) \vdash \mathit{Collapse} & & \text{(F)} \\
(\mathrm{S}^-), \mathit{Collapse}^\Diamond, \exists xx \Box \forall y (y \prec^\Diamond xx) \vdash \bot & & \mathsf{FPFO} \\
(\mathrm{S}^-), \mathit{Collapse}^\Diamond \vdash \forall xx \neg \Box \forall x (x \prec^\Diamond xx) & & \mathsf{FPFO} \\
\Box(\mathrm{S}^-), \Box \mathit{Collapse}^\Diamond \vdash \mathit{Relativism} & & \text{Nec} \\
(\mathrm{S}^-), \mathit{Collapse}^\Diamond \vdash \mathit{Relativism} & & \text{(4)}
\end{array}
$$

This completes the presentation of the arguments from paradox. The relativist is able to give schematic and modal formulations of arguments that, in a logic that should be acceptable to both sides, permits her to derive the schematic and modal formulations of her view from two premisses, Liberalness and Atomic Stability. Let us now turn to assess the two premisses.

3 Liberalness

3.1 Drawing a line

In both its schematic and modal formulations the Liberalness premiss gives voice to a maximally liberal attitude towards set formation among sets: any sets *can* form a set. It is easy to lose sight of just how natural this view is. Over a century's awareness of the paradoxes amongst the mathematical and philosophical community has made us instinctively wary of principles resembling the inconsistent Naïve Comprehension schema. Sometimes, of course, like in the case of Collapse, the suspicion is well founded. But a maximally liberal answer to the formation question need not commit us to such an impossible principle. Liberalness, as articulated schematically or in the modalisation of Collapse, does not lead to contradictions in the accompanying logics. But once a maximally liberal answer to the formation question is seen not to be ruled out on logical grounds, it is hard to see on what other grounds it might be ruled out. What in the world could prevent some sets from being able to form a set?[29]

The distinction between *in fact* forming and *being able* to form a set—in the relevant, non-circumstantial sense—is crucial here. To repeat the familiar point, standard set theory tells us that some sets—such as all the sets, or the von Neumann ordinals—fail to in fact form a set: granted that the present universe $V_\theta[U]$ is closed under the axioms of set theory, no member of that universe has every set in $V_\theta[U]$ or every ordinal in $V_\theta[U]$ as an element. But this is not to say that they cannot form a set. It remains open that all ordinals in the initial universe or all the sets in the initial universe are elements of a set in a more inclusive universe: the *sets* θ and $V_\theta[U]$ are members of the larger universe, $V_{\theta_1}[U]$. But—at least granted Atomic Stability—this distinction collapses for the absolutist. Relative to the absolute perspective $\boldsymbol{p}$, the Invariance thesis tells us that there is no distinction between forming a set relative to $\boldsymbol{p}$, and forming a set relative to another perspective $\boldsymbol{p}'$; the Flattening thesis tells us that there is no distinction between the non-modal condition $\exists y \forall z(z \in y \leftrightarrow z \prec xx)$ articulating that xx do form a set; and its modalisation, articulating that they can form a set. But since some sets in fact do not form a set, the absolutist is forced to conclude moreover that such sets are *unable* to form a set.

Given the initial plausibility of the maximally liberal attitude to set formation this is surprising. And it leaves the absolutist to face a difficult question: what could it be about the world that permits some things to potentially form a set and disqualifies some other things from doing the same? To put the matter more loosely, the absolutist is committed to drawing a line. On one side lie those fortunate 'pluralities' that are able to form sets and on the other those unfortunate 'pluralities' that are unable to form sets. But the commitment to such a line brings with it an explanatory burden. Where does the line fall? And why does it fall where it does?

Some test cases will help to clarify what is being asked for here. The rank of a set may be defined recursively as usual. An ordinal α is said to be the rank of a set if α is the least ordinal greater than the rank of any of its elements.[30] Consider first the

[29]Linnebo (2010) poses the question in much this way and defends a maximally liberal answer.
[30]This relies on the well-foundedness of the element-set relation $\in$ as ensured by Foundation.

pure sets of finite rank $n < \omega$, where ω is the order type of the finite ordinals: $0, 1, 2, \ldots$. These sets comprise the pure hereditarily finite sets, those that can be reached from the empty set by finitely many applications of the operation of adding an element to a set: $x, y \mapsto x \cup \{y\}$. Which side of the line do these fall? What about all those pure sets of accessible rank $\alpha < \theta_0$, where θ_0 is the first (strongly) inaccessible ordinal? These sets comprise those which can be reached from the set of all hereditarily finite sets by any set-theoretic operations: by forming pair-sets subsets, taking powersets, forming unions, or application of replacement. Do they form a set? More exigently, what makes these sets falls on one side of the line rather than other?

As mentioned above, logic does not settle in the absence of absolutism that a line must be drawn at all. In combination with absolutism, plural logic (and Atomic Stability) do ensure that a line must be drawn. But logical consideration tell us almost nothing about *where* the line falls. This combination of assumptions do tell us, for example, that the non-self-membered sets fall the wrong side of the line. But what about our first test case, the pure sets of finite rank? Logic is silent. For all that is logically the case, the sets of finite rank might be all the non-self-membered sets. There is no serious doubt of the logical consistency of the iterative hierarchy's stopping short of ω, or the consistency of its far exceeding θ_0, or that of its achieving any extent in between.

Logic does not determine which sets falls on the right side of the line and which do not, let alone explain why the line gets drawn where it does. The question remains: why is the line drawn where it is? In particular, why are some sets unable to form a set?

3.2 ZFCU

The question appears to be clearly set-theoretic. The obvious place to look for an answer, therefore, is to our best theories of sets, ZFCU and its various relatives and extensions. Such systems immediately settle our first test case in the positive. ZFCU (and indeed its counterpart theory dealing only with pure sets, ZFC) proves the existence of a set V_ω with the pure sets of finite rank as its elements. This provides a mathematical explanation of why the pure sets of finite rank form a set. The axioms of ZFCU, however, are silent on the question of whether or not the pure sets of accessible rank form a set. But this question is settled with the addition of further large cardinal axioms. Similarly, the absolutist might hope that standard set theory may deliver an explanation of why certain sets are *unable* to form a set. For, he may claim, set theory also places negative constraints on set formation. For instance, he may note, it refutes the existence of a universal set. The following sentence is a theorem of ZFCU, and indeed of any subsystem with the Separation axiom.

$$\neg\exists x\forall y(y \in x)$$

Since this is a theorem, he continues, it holds in any model of set theory, and thus the things comprising everything cannot form a set. In this way the absolutist might hope that he has the very beginnings of an account of where the line must be drawn. One that can be elaborated by further set-theoretic investigation.

But any hopes that the line might be drawn and explained by an appeal to standard set theory in this way are quickly extinguished. Notice first that contrary to the absolutist's claim, the theorem in question places *no* limitations on which things are such that *they* can form a set. His argument supports his conclusion: 'the things comprising everything cannot form a set' only if 'cannot' takes wider scope than 'the things'. For given any standard model, that the theorem in question holds in the model, tells us only that those things in the universe of the model—those things that are 'everything' from the perspective of the model—are not the elements of any set in this universe. This leaves open the possibility that *these things* are the elements of a set in the universe of a larger model. (Of course, since the theorem also holds in the larger model, the members of the *larger model's* universe are not the element of a set in the larger domain.)

And in fact this situation is characteristic of set theory at large. Neither ZFCU nor its familiar extensions place any limitations on which things can form a set. While the axioms of set theory, when given a plural or second-order formulation,[31] determine the shape of the set theoretic hierarchy, they fail to determine its extent. Other mathematical theories, such as Peano arithmetic are categorically axiomatised by the plural or second-

[31] The singular axioms fails to determine either the shape or extent of the hierarchy for the familiar Lowenheim-Skolem reasons.

order axioms: their models are uniquely determined up to isomorphism. But standard set theory has not been given a categorical axiomatisation. Even relative to a fixed urelement base, ZFCU achieves only quasi-categoricity. *Any* inaccessible rank $V_\theta[U]$ satisfies the axioms:[32] the axioms admit of larger and larger models.

The additional axioms set theorists propose and investigate place *lower* bounds on the size of the hierarchy, such large cardinal hypotheses succeed in 'raising the ante' as Parson's puts it, but they never place upper bounds on the possible extent of the hierarchy. (1977, p. 523.) Consequently, set-theoretic investigation of the sort large cardinal theorists engage in, generates more and more positive answers to set formation questions of the form 'Can such-and-such objects form a set?'. For example, accepting any large cardinal axiom, enables us to settle our second test case in the positive. But such axioms never generate negative answer to such questions. The absolutist consequently cannot look to ZFCU or its extensions to explain why some things are unable to form a set.

The point is not that it would be hard to achieve a categorical axiomatisation of ZFCU. On the contrary, this is readily achieved by adjoining this theory with the negation of the weakest large cardinal axiom. The point is that set theorists do not do this. It runs contrary to what Parsons describes as 'a general maxim in set theory' that 'any set theory which we can formulate can plausibly be extended by assuming that there is a set that is a (standard) model of it' (1974, p. 218). And indeed this maxim is only the beginning. Set theorists looking for new large cardinal axioms go *much* further than positing the existence of standard models for the currently accepted as axioms. As Gödel puts it:

> there. . . exist unexplored series of axioms which are analytic in the sense that they only explicate the content of the concepts occurring in them, e.g., the axioms of infinity in set theory, which assert the existence of sets of greater and greater cardinality or of higher and higher transfinite types and which only explicate the content of the general concept of set (1972, p. 306)

To seek to cut off the set-theoretic hierarchy at any inaccessible rank by positing the

[32] Zermelo (1930).

negation of a large cardinal axioms runs directly against set-theoretic practice.

3.3 Limitation of size

If set theory itself does not draw the line that the absolutist who denies Liberalness is committed to, he might look instead to the ideas that have played a role in its historical development. One such thought is that some sort of 'limitation of size' constraint has a role to play in determining whether some sets do or do not form a set. For example, Azriel Levy describes it as 'the guiding principle' behind standard set theory that the Naïve Comprehension schema is restricted to instances that 'assert the existence of sets which are not too 'big' compared to sets which we already have' (Fraenkel, Bar-Hillel and Levy 1973, p. 32)[33]

Accordingly the absolutist might seek to draw a line on the basis of size.

Limitation of Size. Some sets form a set only if they are not 'too many'.

But what is the relevant standard: how many are 'too many'? If the line is to be drawn on this basis we need to see that the relevant standard settles where the line falls in a well-motivated way.

Two options are apparent. The first option is to look for an absolute standard. For example, we might decree that some sets are 'too many' just in case they are greater in number than (a) forty two, say, or alternatively, (b) if they exceed the forty-third smallest infinite cardinal, $\aleph_{42}$; or perhaps, (c) the forty-third smallest inaccessible cardinal θ_{42}. Of course the cardinality bound should not invoke sets if the explanation of where the line is to be drawn is to avoid circularity. And in these cases this can be done by drawing on higher-order resources.[34] But the hopes for *motivating* such a standard seem vanishingly small. Any absolute bound would be irredeemably arbitrary.

The second option is to look for a relative standard. For example it may be maintained that some sets are 'too many' just in case they are as numerous as the ordinals. Such standards are less obviously unacceptably arbitrary.[35]

[33] See Hallett (1984, ch. 5) for a historical overview, and criticism.
[34] See Shapiro (1991, §5.1.2).
[35] See however Linnebo (2010) for an argument that they are circular.

All the same, regardless of whether or not explanations in this vein can be sustained, a relative standard of Limitation of Size offers no advance to the absolutist on what was offered to him by standard set theory. For this sort of Limitation of Size principle has already been assimilated into standard set theory. It is a theorem of ZFCU (when given a plural formulation) that any things form a set only if they are no more numerous than the ordinals. Granted that standard set theory fails to provide the explanation the absolutist needs, so too does the vacuous extension of this theory by the relative-standard Limitation of Size principle.

Consequently, Limitation of Size fails to draw a well-motivated line. An absolute standard of being 'too many' succeeds for the first time is drawing a firm line but is ill motivated. A relative standard offers no advance over standard set theory.

3.4 The iterative conception

A second influential picture behind ZFCU and its extensions is the iterative conception of set, popularised by Boolos (1971).[36]

We can think of sets as being formed in a well-ordered sequence of stages. At the very first stage, there are no sets, and none are formed. At each subsequent stage, all sets of sets formed at earlier stages will be formed. At the second stage, $\varnothing$ will be formed; at the third, $\varnothing$ and $\{\varnothing\}$ will be formed; at the fourth, $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$ and $\{\varnothing, \{\varnothing\}\}$ will be formed; similarly for the fifth, the sixth, and so on. The set formation process is iterated through all the stages

Iterative Conception Some sets form a set just in case they are all available at some stage in the sequence.

Problematic sets like all the non-self-membered sets, or all von Neumann ordinals, do not form a set because there is no stage by which all of their members have been formed. At each stage, new non-self-membered sets, and new ordinals are formed. This conception does succeed in drawing the line the absolutist is committed to—but only relative to the sequence of stages. If, as is usually assumed, the sequence of stages is

[36]It was presented earlier by Shoenfield (1967).

longer than the sequence of finite ordinals, then the set of sets of finite rank V_ω is formed at the first limit stage. Whether or not the sets of accessible rank form a set depends on whether the sequence of stages extends far enough.

Now, as we shall see in the next chapter, I think there is a great deal to be said for this conception when suitably interpreted. But it is unclear what the absolutist is to make of it. Few absolutists will want to take seriously the metaphor of sets being formed in a temporal process.[37] Boolos's view is that neither the tense in this informal sketch of the iterative conception, nor the metaphor of set 'formation', should be taken at all seriously. The informal sketch is just a colourful story or 'narrative convention'[38] that ultimately gives way to a formal axiomatisation of the iterative conception, in the form of a 'stage theory'.[39] This raises an initial worry. If stage theory is a mere narrative convention, why think that it can play a serious explanatory role?

But regardless of whether this worry can be met there is a more serious problem. The iterative conception succeeds in drawing a well-motivated line only if a well-motivated account of when the stages give out can be extracted from the antecedent theory of stages. But this *explanandum* would appear to be every bit as difficult to account for as the original one. For 'stage' would seem to have just as good a claim to be indefinitely extensible as 'set' does. To decide to cut off the stages at any given order-type seems gratuitously arbitrary. Why could we not iterate the process further?

3.5 A brute fact?

This exhausts all of the obvious places the absolutist might look to help him draw a line. Neither logic, nor standard set theory and its extensions, nor the motivations behind such theories seems to provide the explanation the absolutist requires.

Of course, explanations must end somewhere. Might the absolutist not claim that it is simply a brute fact about the world, not open to mathematical explanation, where

[37] Wang (1974) is a notable exception; but views of this sort face powerful objections; see chapter 5.

[38] Boolos (1989), p. 90, attributes the thought that the iterative conception is a narrative convention, describing of the set-theoretic hierarchy, according to one salient aspect of its structure—rank by rank—to Dan Leary.

[39] Boolos gives two axiomatisations of stage theory. See Boolos (1971), pp. 20–22, and Boolos (1989), p. 91.

the troublesome line is to be drawn. We do not expect our best physical theories, to deliver precise verdicts on the boundaries between filled and unfilled regions of spacetime, settling for instance, whether or not a particular unobserved spacetime region R_0 contains matter. Likewise, the absolutist might claim, we should not expect our best mathematical theories to draw a particular line between the things that are and are not able to form a set, determining, for example, whether or not some particular things xx_0 form a set. What is required of the physical theory is that it should deliver general predictions that match up with and explain the particular physical facts that we do observe, delivering, for example, general claims about the distribution of matter across spacetime (without needing to settle the particular question of whether R_0 is filled). Likewise, he may claim, all that we should expect from mathematical theories like ZFCU, is that they should yield general principles about sets, as typified in the theorems of set theory (without needing to settle whether xx_0 can form a set). But just because our best theory fails to deliver an answer does not mean that there is no fact of the matter—in either case.

The trouble with such a stance is that it seems to render quite mysterious how we might come to know such brute facts about the mathematical universe. The absolutist who takes this tack may be tempted to respond that he is not a verificationist. There are all sorts of facts that he does not know and could not find out. And in some cases there may be physical reasons that put the matter distribution facts beyond our epistemic reach: if R_0 is a region of spacetime outside our past light cone, and thus causally inaccessible from our location, for instance, we may be simply unable to find out whether or not it is filled. But this does not mean that there is no fact of the matter.

The mathematical and physical case, however, are importantly different. And this difference is manifested in how the scientific and mathematical communities go about investigating their respective subject matters. In the case of brute facts about the physical world, we are able to come to know at least some of them through perception. On this basis we can set about systematising and generalising from our perceptual knowledge in the familiar way. In particular we can set about revising and refining our theories when they are found not to meet the data. But in the mathematical case, very few philosophers of mathematics have been content to posit some perception-like faculty through which we might come to learn particular mathematical facts.[40] Without such a faculty we are unable to come to know *any* of the brute facts in this way. It seems, therefore, that unlike the physical case, to posit brute facts in the mathematical case is to posit facts that float wholly free of our usual routes to mathematical beliefs. But granted that whether xx_0 form a set is a simply a brute fact, and that mathematical theorising has no epistemic link to such facts, what reason do we have to think that standard set theory gets it right? For all we know, it might be a brute fact that the sets of finite rank fail to form a set after all; just as it might be a brute fact that R_0 is unfilled. Contraposing, granted that mathematical practice *is* along the right lines, and set theory *has* furnished us with a wealth of knowledge about particular matters of mathematical fact—the existence of V_ω amongst them—the physical analogy appealed to by the absolutist cannot be along the right lines. Unexplained arbitrariness is unacceptable in mathematics in a way that is not unacceptable in the physical sciences.

4 Atomic Stability

4.1 Soft indefinite extensibility

The emergence of a second premiss in our regimentation of the argument from paradox provides a second potential way out for the absolutist. *Prima facie*, at least, the various formulations of Atomic Stability do not have anything like the intuitive pull of Liberalness. Granted that the relativist postulates shifts in the extensions of expressions with shifts in perspective, why should we expect even the limited constancy that Atomic Stability posits?

And, of course, rejecting the second key premiss permits the absolutist to accept the first. For instance he can accept a metalinguistic formulation of Liberalness, dropping the restriction on the quantifier he contends is vacuous.

Liberalness (*Metalinguistic*). For any initial sets some new set has them as elements.

[40] Gödel's supplement to his 1947 paper is a famous exception. See Gödel (1983). See also Maddy (1980).

$$\forall xx\big(B_1 xx \to \exists y(\beta_2 y \wedge \forall z(z \in_2 y \leftrightarrow z \prec_2 xx))\big)$$

And he can likewise accept the schematic and modal formulation of Liberalness. On this basis the absolutist may hope to avoid drawing the troublesome line. He accepts a 'soft' version of indefinite extensibility. Given any initial extension of 'set' (or its formal analogue: β) we can come to a more liberal extension containing an object not in the initial extension, leading to correlative expansions in the extension of 'element' (or $\in$). The only difference between this view and the 'hard' version upheld by the relativist is that the new set is not a new object. Such a view still upholds the 'positive' stability of the predicates (S^+) but rejects the 'negative' stability (S^-) deployed in the relativist's argument.

(S^+) $\forall_1 x(\beta_1 x \to \beta_2 x)$

(S^-) $\forall_1 x(\beta_2 x \to \beta_1 x)$

The picture that emerges is something like Zermelo's (1930) account of the set-theoretic hierarchy. There is an unbounded sequence of larger and larger hierarchies of pure sets V_{θ_0}, V_{θ_1}, ... Sets$_\alpha$ that do not form a set$_\alpha$, relative to the α-th hierarchy, such as instance the ordinals$_\alpha$, or all the sets$_\alpha$ becomes perfectly good sets$_{\alpha+1}$ within the next hierarchy, namely θ_α and V_{θ_α}. But this all happens *within* the ultimate domain. In coming to new sets, we do not arrive at new objects. Each larger and larger inaccessible rank V_{θ_α} is properly contained within the ultimate domain $\boldsymbol{M}$.

This calls for some account of *how* 'set' and 'element' (or their formal analogues β and $\in$) shift their extension. But here a similar range of options will be available to the absolutist as is available to the relativist. In particular it may be maintained that these predicates are context-sensitive, or that the shifts in extension are due to a shift in interpretation of the lexicon. Williamson (1998b) has briefly sketched a view of the second sort. The lesson to be drawn from indefinite extensibility is that 'given any reasonable assignment of meaning to the word "set" we can assign it a more inclusive meaning while feeling that we are going on in the same way' (p. 20).[41] As with interpre-

[41] The principal focus of this paper is the Liar paradox. Williamson develops a response based on the idea that true shifts its extension in virtue of subtle shifts in meaning. He briefly

tational relativism, not any arbitrary reinterpretation will be acceptable. In the second case, an account needs to be given of what extensions of 'set' are admissible, containing only objects that warrant the label. Still, if the relativist can give an acceptable account there seems no reason to doubt that the absolutist can follow suit.

However, despite its initial appeal, the attempt to reject Atomic Stability leads to two severe problems.

4.2 Urelements

The first problem is that this means of rejecting Atomic Stability leads to trouble with urelements. In order to integrate set theory with other domains of enquiry we need to focus on set theory with urelements, ZFCU, rather than merely the study of pure sets, ZFC. For instance if we require functions mapping physical objects to their mass, say, or the extension of the predicate 'person', we require impure sets: sets that contain non-sets in their transitive closure. Standardly 'urelement' is defined simply as non-set: $\neg ßx$.

But, for the proponent of soft indefinite extensibility, this means that every future set lurks hidden among the urelements. For something is an urelement$_1$ if and only if it is not a set$_1$; or an urelement$_2$ if and only if not a set$_2$; and so on. Consequently, any 'proper' set$_2$—any set$_2$ that is not a set$_1$—is therefore an urelement$_1$.

Stewart Shapiro observes that this has some unsettling consequences. 'Cardinality considerations (after a fashion) entail that most of the ordinals$_2$ cannot be pure sets$_1$.' But since the sets$_2$ are urelements$_1$, '[i]t follows that... Vann McGee's urelement set axiom, stating that the urelements form a set, fails and fails badly in any language which has a set theory like ZFC and a quantifier ranging over absolutely everything.' (2003, p. 477)

The failure of the Urelement Set axiom for sets$_1$ can be seen from the perspective of the standard set theory for sets$_2$. For sets$_1$, the axiom is as follows.

Urelement Set Axiom. Some set$_1$ has every urelement$_1$ as a member.

discusses this idea in application to set and indefinitely extensibility, towards the end his paper, pp. 19-20. See Shapiro (2003, pp. 473–478) for discussion.

(S) $\exists y(\text{ß}_1 y \wedge \forall z(\neg \text{ß}_1 z \rightarrow z \in_1 x))$

In line with soft indefinite extensibility suppose that between the first and second perspective the extension of ß and $\in$ expands so that we still have positive stability for these predicates

(S$^+$) $\forall x(\text{ß}_1 x \rightarrow \text{ß}_2 x)$

$\forall x \forall y(x \in_1 y \rightarrow x \in_2 y)$

Suppose moreover that standard set theory with urelements holds good of both sets$_1$ and sets$_2$; so that the ZFCU axioms hold when ß and $\in$ are replaced by ß_1 and $\in_1$ or by ß_2 and $\in_2$, under the intended interpretation of these predicates. Then the Urelement Set axiom is incompatible with the Liberalness principle that the rejection of Atomic Stability aimed to save. For the Urelement Set axiom ensures that $\{x \mid \neg \text{ß}_1 x\}$ is a set$_1$ with every urelement$_1$ as an element$_1$. Consequently (by Positive Stability) this is also a set$_2$ with every urelement$_1$ as an element$_2$. But Liberalness ensures that we have a set$_2$ comprising every set$_1$ as an element$_2$, $\{x \mid \text{ß}_1 x\}$. Applying Union we obtain a universal set$_2$ $\{x \mid \neg \text{ß}_1 x \vee \text{ß}_1 x\}$, and this leads to a contradiction via Separation and the Russell argument.

But actually as well as giving up what many absolutists have seen as an attractive *addition* to ZFCU, to follow this strategy the absolutist must cut into standard set theory for urelements itself, when this is given a higher-order or plural formulation. For consider the function, encoded by a higher order or plural term denotation, $F\colon \{x\} \mapsto x$, that maps each singleton$_1$ to its sole element$_1$ (and, say, maps everything else to the empty set$_2$). The Pairing axiom for sets$_1$, ensures that any thing whatsoever has a singleton$_1$.

Pairing$_1$ $\forall x_1 \forall x_2 \exists y(\text{ß}_1 y \wedge \forall z(z \in_1 y \leftrightarrow z = x_1 \vee z = x_2)$

Consequently this function is surjective onto the entire universe, even when restricted only to sets$_1$. But since Liberalness provides a set$_2$ $\{x \mid \text{ß}_1 x\}$ of every set$_1$, the image of this set$_2$ under the function F, $\{F(x) \mid x \in \{x \mid \text{ß}_1 x\}\}$ is a set$_2$ by Replacement. And we have once again reached a universal set$_2$, and a contradiction.

The upshot is that to follow this response, the absolutist must restrict the Pairing axiom. This is a serious concern, for while changes to the higher reaches of set theory may have little impact outside set theory, restrictions to Pairing cut out sets at the lowest ranks. And this may have severe repercussions outside pure mathematics. For, as is widely recognised it is our ability to form sets of urelements, with operations like Pairing, that ensures we have set-theoretic structures rich enough for the needs of applied mathematics.[42]

The absolutist may rejoin that it is only the presence of future members of the extension of 'set' amongst the urelements that threatens Pairing. And these are *not* required for ordinary applied mathematics. What is needed then is a restriction that prohibits proper sets_2, proper sets_3 and so on from forming singletons_1 but does not leave us without the pair sets needed for applied mathematics. The question remains how? The obvious answer is to have a predicate for 'ultimate urelement' which applies, relative to any perspective, to those things which never admissibly enter the extension of 'set'. Unfortunately, the absolutist cannot appeal to such a predicate.

4.3 Revenge

The availability of an ultimate urelement predicate leads directly to the first of two revenge problems that threaten to undermine the soft indefinite extensibility strategy.

The success of the soft indefinite extensibility response to the arguments from paradox hinges on there being an unbounded sequence of more and more liberal extensions for set. If we could collect all of the future sets lurking amongst the urelements into a single, 'ultimate' extension of 'set', applying to sets_0, sets_1, ..., the absolutist would no longer be able to resist Negative Stability. For relative to any other perspective, any set_α in an admissible extension of 'set' would *ipso facto* be a member of such an ultimate extension. And granted Stability, the absolutist's only remaining option is to give up Liberalness. If we can come to an ultimate extension of 'set', the soft indefinite extensibility strategy fails.

The trouble with an ultimate urelement predicate is that it leads directly to the

[42]See, for instance, McGee (1997), Uzquiano (2006) and Williamson (2006, p. 386).

troublesome ultimate set predicate. Granted that the ultimate urelement predicate performs its intended function and applies only to things that are not sets_0, not sets_1, not sets_2, . . . , its negation functions as an ultimate set predicate: it applies to things that are either sets $_0$ or sets_1 or sets_2 or . . . . To follow the soft indefinite extensibility response to the argument from paradox, therefore, the absolutist must do without an ultimate urelement predicate, and accordingly do without a maximally liberal formulation of the Pairing axiom.

But setting the Pairing issue to one side, we might wonder what is to stop us gathering all of the eventual sets within a single extension. If the soft indefinite extensibility strategy is to be a robust response to the argument from paradox, there needs to be a similarly robust reason why we cannot thus come to such ultimate extensions. Williamson in the case of 'true' points to the unpredictability of shifts in meaning. He contends that we are unable to anticipate future admissible interpretations of the predicate on the basis of our present our current understanding. Consequently, we are unable to gather the future admissible extension into a single ultimate extension. (1998a, pp. 16–17)

An initial worry with this explanation is that it threatens to over-generate. The absolutist wants to maintain that we *can* come to an ultimate extension for logical expressions like 'thing', the identity predicate = and the quantifier $\forall v$. If admissible shifts in extension are unpredictable, in a way that prevents us from pinning down an ultimate extension for $\in$, why suppose that we can pin down an ultimate extension for the logical expressions? Perhaps the response is that *logical* expressions do not shift their meaning in an unpredictable way. But I wonder whether the logical/mathematical distinction does the work required of it here. For despite its differences from logical expressions, the admissible interpretations of β do not appear to be harder to anticipate. On the contrary, we have a very full account of what the more extensive extensions will be like.

Moreover, to the extent that the pluralist's inter-perspective modal operator is intelligible, predicates playing the problematic role can be obtained by fiat. For we may define the predicates $\textit{ß}$ and $\boldsymbol{\varepsilon}$ as follows:

$$\textit{ß}v =_{\mathrm{df}} \Diamond \textit{ß}v \qquad u \mathrel{\boldsymbol{\varepsilon}} v =_{\mathrm{df}} \Diamond u \in v$$

This allows us to eliminate any appeal to Atomic Stability in the modal version of the argument from paradox. Write $\phi^{\star}$ for the result of replacing each subsentence in ϕ of the form $\textit{ß}v$ and $u \in v$ with $\textit{ß}v$ and $u \mathrel{\boldsymbol{\varepsilon}} v$. Since prefixing such subsentences with modal operators automatically ensures their stability, we can prove the following version of the Flattening thesis.

Flattening (*without Atomic Stability*)

(F*) $\exists xx \Box \forall y(y \prec^{\Diamond} xx) \vdash \phi^{\star} \leftrightarrow \phi^{\Diamond}$

Moreover, unlike modalisation, such a substitution does nothing to relieve the inconsistency of *Collapse*: *Collapse** leads to a contradiction by the standard argument, using the instance of Plural Comprehension for $\textit{ß}x \wedge \neg x \mathrel{\boldsymbol{\varepsilon}} x$. Consequently, drawing on the revised flattening thesis, the relativist can modify her earlier modal argument. Relativism is derived by modal reasoning just from Liberalness as articulated in *Collapse*$^{\Diamond}$.

$$\begin{aligned}
\textit{Collapse}^{\Diamond}, \exists xx \Box \forall y(y \prec^{\Diamond} xx) &\vdash \textit{Collapse}^{\star} \\
\textit{Collapse}^{\Diamond}, \exists xx \Box \forall y(y \prec^{\Diamond} xx) &\vdash \bot \\
\textit{Collapse}^{\Diamond} &\vdash \forall xx \neg \Box \forall x(x \prec^{\Diamond} xx) \\
\Box \textit{Collapse}^{\Diamond} &\vdash \textit{Relativism} \\
\textit{Collapse}^{\Diamond} &\vdash \textit{Relativism}
\end{aligned}$$

Granted the logic for the modal language, the argument from paradox need rely only on Liberalness. The only remaining option for the absolutist is to reject this assumption.

The first revenge problem attempts to bring the sets$_0$, sets$_1$, sets$_2$, and so on, into the extension of a single predicate for 'set'. The second does not rely on such a predicate. The question of liberalness as formulated so far has focused on the question of which *sets* are able to form a set. This is a special case of the more general question: which *things* are able to form a set? But the motivations for a maximally liberal answer to the specific question carry over directly to motivating a maximally liberal answer to the second. It is no easier to draw a line between which things do and do not form a set than it is to draw a line between which sets do and do not form a set. And as in the previous section, we should be wary of an approach which posits unexplained arbitrariness in this setting.

A truly maximally liberal attitude to set formation contends not merely that any sets can form a set, but that any things can form a set. In the case of a schematic formulation, Liberalness is fully articulated by dropping the restriction to those things which are sets from the earlier sequent.

Liberalness (Schema) $\varsigma\colon E\tau\tau \Rightarrow \varsigma'\colon \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec \tau\tau))$

But when Liberalness is thus articulated, the soft indefinite extensibility strategy response no longer helps. This principle permits us to give a schematic derivation of relativism without appeal to Atomic Stability.

For consider $rr^* =_{\text{df}} \langle \lambda_{\varsigma'} x.x \notin x\rangle$. It is straightforward to adapt the sequent version of the Russell argument on page 123 in order to derive the following sequent.

$$\varsigma'\colon \exists y(\text{ß}y \wedge \forall z(z \in y \leftrightarrow z \prec rr^*)) \Rightarrow$$

In conjunction with Liberalness, we obtain $\varsigma\colon Err^* \Rightarrow$. Finally, on the basis of the sequent rules for logical expressions displayed in table 4.5, it is straightforward to derive this sequent.

Descent $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E\tau\tau \Rightarrow \varsigma\colon E\tau\tau$

(See the appendix.) We then reason as follows, obtaining the initial right sequent from the sequent version of Plural Comprehension.

$$\dfrac{\dfrac{\dfrac{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon Err^* \Rightarrow \varsigma\colon Err^* \qquad \Rightarrow \varsigma'\colon Err^*}{\varsigma'\colon \forall x(x \prec mm_\varsigma) \Rightarrow \varsigma\colon Err^*} \qquad \varsigma\colon Err^* \Rightarrow}{\varsigma'\colon \forall x(x \prec mm_\varsigma) \Rightarrow}}{\Rightarrow \neg\varsigma'\colon \forall x(x \prec mm_\varsigma)}$$

We can argue directly from the relevant instance of the more general formulation of Liberalness to the relevant instance of the relativist's schema.

Consequently the soft indefinite extensibility strategy does not in the end offer the absolutist an alternative to taking a less than maximally liberal attitude to set formation. If we can come to an ultimate set predicate—as we can if the pluralist's modal operator is available to us, this strategy fails altogether, the Liberalness principles stating that any sets can form a set which the absolutist sought to save by admitting shifts in the extension of 'set' yields Relativism. But even if we cannot come to an ultimate sense of 'set', soft indefinite extensibility falls short of endorsing a maximally liberal attitude to which *things* can form a set. In the end, the absolutist has no option but to draw a line. But as we saw above there seems to be no principled place to draw it.

5 Appendix: Proofs

5.1 Schematic proofs

Invariance is proved via three lemmas, Ascent, Descent, and Atomic Invariance.

Invariance. The following are derivable from (S^+) and (S^-).

(I^+) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi \Rightarrow \varsigma'\colon \phi$

(I^-) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \phi \Rightarrow \varsigma\colon \phi$

Ascent. The following are both instances of (S^+)

(1) $\varsigma\colon E\tau \Rightarrow \varsigma'\colon E\tau$

(2) $\varsigma\colon E\tau\tau \Rightarrow \varsigma'\colon E\tau\tau$

Descent The following are derivable.

(3) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E\tau \Rightarrow \varsigma\colon E\tau$

(4) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E\tau\tau \Rightarrow \varsigma\colon E\tau\tau$

Proof. For (3), apply cut to the following two derivations.

$$\dfrac{\dfrac{}{\varsigma'\colon \tau \prec mm_\varsigma \;\Rightarrow\; \varsigma'\colon \tau \prec mm_\varsigma}\,\text{REP} \qquad \dfrac{}{\varsigma'\colon E\tau \;\Rightarrow\; \varsigma'\colon E\tau}\,\text{REP}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E\tau \;\Rightarrow\; \varsigma'\colon \tau \prec mm_\varsigma}\,\text{L}\forall v$$

$$\dfrac{\dfrac{}{\varsigma'\colon \tau \prec mm_\varsigma, \varsigma\colon Emm_\varsigma \;\Rightarrow\; \varsigma\colon E\tau}\,\text{L}Evv \qquad \dfrac{}{\Rightarrow\; \varsigma\colon Emm_\varsigma}\,\text{PC}}{\varsigma'\colon \tau \prec mm_\varsigma \;\Rightarrow\; \varsigma\colon E\tau}\,\text{CUT}$$

For (4), reason as follows.

$$
\dfrac{\dfrac{\begin{array}{c}(3)\\ \vdots \\ \varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon Ex \ \Rightarrow\ \varsigma\colon Ex\end{array} \qquad \dfrac{}{\varsigma'\colon E\tau\tau, \upsilon\colon x \prec \tau\tau \ \Rightarrow\ \varsigma'\colon Ex}\,\textsc{l}E\upsilon\upsilon}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E\tau\tau, \upsilon\colon x \prec \tau\tau \ \Rightarrow\ \varsigma\colon Ex}\,\textsc{cut}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E\tau\tau \ \Rightarrow\ \varsigma\colon E\tau\tau}\,\textsc{r}E\upsilon\upsilon
$$

In this derivation υ is a perspective variable not free in ς □

Atomic Invariance The following are derivable from (S^+) and (S^-).

(5) $\varsigma\colon \Phi \Rightarrow \varsigma'\colon \Phi$

(6) $\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \Phi \Rightarrow \varsigma\colon \Phi$

Proof. The following is readily derivable from the rules for singular and plural identity and conjunction, for Φ atomic:

(7) $\varrho\colon \Phi \Rightarrow \varrho\colon E_\Phi$

For (5), we then reason as follows

$$
\dfrac{\varsigma\colon \Phi, \varsigma\colon E_\Phi \ \Rightarrow\ \varsigma'\colon \Phi \qquad \begin{array}{c}(7)\\ \vdots \\ \varsigma\colon \Phi \ \Rightarrow\ \varsigma\colon E_\Phi\end{array}}{\varsigma\colon \Phi \ \Rightarrow\ \varsigma'\colon \Phi}\,\textsc{cut}
$$

Proof of (6)

$$
\dfrac{\dfrac{\varsigma'\colon \Phi, \varsigma\colon E_\Phi \ \Rightarrow\ \varsigma\colon \Phi \qquad \begin{array}{c}(3) \text{ and } (4)\\ \vdots \\ \varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon E_\Phi \ \Rightarrow\ \varsigma\colon E_\Phi\end{array}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \Phi, \varsigma'\colon E_\Phi \ \Rightarrow\ \varsigma\colon \Phi}\,\textsc{cut} \qquad \begin{array}{c}(7)\\ \vdots \\ \varsigma'\colon \Phi \ \Rightarrow\ \varsigma'\colon E_\Phi\end{array}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \Phi \ \Rightarrow\ \varsigma\colon \Phi}\,\textsc{cut}
$$

□

Proof of Invariance. We prove (I^+) and (I^-) simultaneously by induction on the complexity of ϕ. Atomic invariance provides the base case. Here are the proofs of (I^+).

(i) $\phi = \neg\phi_1$.

$$
\dfrac{\dfrac{\begin{array}{c}(\mathrm{IH})\\ \vdots \\ \varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \phi_1 \ \Rightarrow\ \varsigma\colon \phi_1\end{array}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon \phi_1, \varsigma\colon \neg\phi_1 \ \Rightarrow}\,\textsc{l}\neg}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \neg\phi_1 \ \Rightarrow\ \varsigma'\colon \neg\phi_1}\,\textsc{r}\neg
$$

(ii) $\phi = \phi_1 \wedge \phi_2$.

$$\dfrac{\dfrac{\overset{\text{(IH)}}{\vdots}\atop{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi_1 \Rightarrow \varsigma'\colon \phi_1} \qquad \overset{\text{(IH)}}{\vdots}\atop{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi_2 \Rightarrow \varsigma'\colon \phi_2}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi_1, \varsigma\colon \phi_2 \Rightarrow \varsigma'\colon \phi_1 \wedge \phi_2}\ \mathrm{R}\wedge}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi_1 \wedge \phi_2 \Rightarrow \varsigma'\colon \phi_1 \wedge \phi_2}\ \mathrm{L}\wedge$$

(iii) $\phi = \forall v\phi$.

$$\dfrac{\dfrac{\overset{\text{(IH)}}{\vdots}\atop{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \phi_1 \Rightarrow \varsigma'\colon \phi_1} \qquad \overset{(3)}{\vdots}\atop{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma'\colon Ev \Rightarrow \varsigma\colon Ev}}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \forall v\phi_1, \varsigma'\colon Ev \Rightarrow \varsigma'\colon \phi_1}\ \mathrm{L}\forall v}{\varsigma'\colon \forall x(x \prec mm_\varsigma), \varsigma\colon \forall v\phi_1 \Rightarrow \varsigma'\colon \forall v\phi_1}\ \mathrm{R}\forall v$$

(iv) $\phi = \forall vv\phi$ is proved in the same way replacing (3) with (4).

The cases for (I$^-$) is proved similarly, replacing (3) and (4) with (1) and (2) in the quantifier cases. $\square$

5.2 Modal proofs

Write $\vdash$ for provable in $\mathsf{FPFO} + \mathsf{S5} + \mathrm{RBF} + \mathrm{RBF}_p + \Diamond Ev + \Diamond Evv$. The Flattening theses are proved via two lemmas, Ascent and Modalised Invariance

Flattening

(F) $\exists xx\Box\forall y(y \prec^\Diamond xx), (\mathrm{S}^-), \vdash \phi \leftrightarrow \phi^\Diamond$

(F*) $\exists xx\Box\forall y(y \prec^\Diamond xx) \vdash \phi^\star \leftrightarrow \phi^\Diamond$

Ascent (Modal)

(8) $Exx, \Box\forall y(y \prec^\Diamond xx) \vdash Ez$

(9) $Exx, \Box\forall y(y \prec^\Diamond xx) \vdash Ezz$

(10) $Exx, \Box\forall y(y \prec^\Diamond xx) \vdash E_\Phi$

Proof of Ascent. Start with (8). The following are routinely derivable in K and KB:

(11) $\Box\psi_1 \rightarrow \psi_2 \rightarrow (\Diamond\psi_1 \rightarrow \Diamond\psi_2)$

(12) $\Diamond\Box\phi \rightarrow \phi$

We then reason as follows.

$$\begin{array}{rcll}
\forall y(y \prec^{\Diamond} xx), Ez & \vdash & \exists y(y \prec^{\Diamond} xx \wedge z = y) & \textsf{FPFO} \\
\Box\forall y(y \prec^{\Diamond} xx), \Diamond Ez & \vdash & \Diamond\exists y(y \prec^{\Diamond} xx \wedge z = y) & \text{DT, Nec, (11), MP} \\
Exx, \Box\forall y(y \prec^{\Diamond} xx), \Diamond Ez & \vdash & \exists y(y \prec^{\Diamond} xx \wedge \Diamond z = y) & \text{RBF} \\
Exx, \Box\forall y(y \prec^{\Diamond} xx), \Diamond Ez & \vdash & \exists y(\Diamond z = y) & \textsf{FPFO}
\end{array}$$

Finally, we show that the $\Diamond$ can be dropped from the conclusion.

$$\begin{array}{rcll}
 & \vdash & \Box(Ez \to z = z) & \textsf{FPFO}\text{, Nec} \\
z = y, \Box(Ez \to z = z) & \vdash & \Box(Ey \to z = y) & \textsf{FPFO} \\
z = y & \vdash & \Box(Ey \to z = y) & \text{MP} \\
\Diamond z = y & \vdash & \Diamond\Box(Ey \to z = y) & \text{DT, Nec, (11)} \\
Ey, \Diamond z = y & \vdash & z = y & \text{(12), MP} \\
\exists y(\Diamond z = y) & \vdash & \exists y(z = y) & \textsf{FPFO}
\end{array}$$

For (9), we can show $Exx, \Box\forall yy(yy \prec^{\Diamond} xx) \vdash Ezz$ in the same way replacing RBF with its plural analogue. (9) follows since it is readily verified that

(13) $\Box\forall y(y \prec^{\Diamond} xx) \vdash \Box\forall yy(yy \prec^{\Diamond} xx)$

(10) is then immediate from (8) and (9). $\Box$

Modalised Invariance

(14) $\vdash \phi^{\Diamond} \leftrightarrow \Box\phi^{\Diamond}$

(15) $\vdash \phi^{\Diamond} \leftrightarrow \Diamond\phi^{\Diamond}$

Proof of Modalised Invariance. We rely on the following standard theorems of $\textsf{S5}$

(16) $\neg\Diamond\psi \leftrightarrow \Box\neg\psi$

(17) $\neg\Box\psi \leftrightarrow \Diamond\neg\psi$

(18) $\Box(\psi_1 \wedge \psi_2) \leftrightarrow \Box\psi_1 \wedge \Box\psi_2$

(19) $\Diamond(\psi_1 \wedge \psi_2) \to \Diamond\psi_1 \wedge \Diamond\psi_2$

(20) $\Diamond\psi \leftrightarrow \Box\Diamond\psi$

(21) $\Diamond\psi \leftrightarrow \Diamond\Diamond\psi$

(22) $\Box\psi \leftrightarrow \Box\Box\psi$

(23) $\Box\psi \leftrightarrow \Diamond\Box\psi$

The proof then goes by induction on complexity.

(i) $\phi = \Phi$, atomic; $\phi^{\Diamond} = \Diamond\Phi$. And (14) and (15) are immediate by (20) and (21).

(ii) $\phi = \neg\phi_1$; $\phi^{\Diamond} = \neg\phi_1^{\Diamond}$. Contraposing the induction hypothesis (IH) (15) and (14), we have: $\neg\phi_1^{\Diamond} \leftrightarrow \neg\Diamond\phi_1^{\Diamond}$ and $\neg\phi_1^{\Diamond} \leftrightarrow \neg\Box\phi_1^{\Diamond}$. (14) and (15) follow by (16) and (17).

(iii) $\phi = \phi_1 \wedge \phi_2$; $\phi^{\Diamond} = \phi_1^{\Diamond} \wedge \phi_2^{\Diamond}$. Applying propositional reasoning to the IH, yields $\phi_1^{\Diamond} \wedge \phi_2^{\Diamond} \leftrightarrow \Box\phi_2^{\Diamond} \wedge \Box\phi_2^{\Diamond}$ and $\Diamond\phi^{\Diamond} \wedge \Diamond\phi_2^{\Diamond} \rightarrow \phi_1^{\Diamond} \wedge \phi_2^{\Diamond}$. (14) and the right-to-left direction of (15) follow by (18) and (19). The left-to-right direction of (15) is immediate by T.

(iv) $\phi = \forall v\phi_1; \phi^{\Diamond} = \Box\forall v\phi_1^{\Diamond}$. (14) and (15) are immediate by (22) and (23).

(v) $\phi = \forall vv\phi_1$ follows the same argument as the singular case. $\Box$

Proof of (F) by induction on complexity.

(i) $\phi = \Phi$, atomic; $\phi^{\Diamond} = \Diamond\Phi$. The left-to-right direction is immediate from T. For the other direction, we use Atomic Stability (S^-), then Ascent (10).

$$(\mathrm{S}^-) \vdash \Diamond\Psi \to (E_\Psi \to \Psi) \qquad \text{Prop. logic}$$

$$Exx, \Box\forall y(y \prec^{\Diamond} xx), (\mathrm{S}^-) \vdash \Diamond\Psi \to \Psi \qquad \text{(10), Prop. logic}$$

(ii) The case for connectives are immediate.

(iii) $\phi = \forall v\phi_1; \phi^{\Diamond} = \Box\forall v\phi_1^{\Diamond}$. The right-to-left direction is immediate from IH and T. For the other direction, the IH and Modalised Invariance yields

$$Exx, \Box\forall y(y \prec^{\Diamond} xx), (\mathrm{S}^-), \phi_1 \vdash v \prec^{\Diamond} xx \to \Box\phi_1^{\Diamond} \qquad \text{(14), }\mathsf{FPFO}$$

$$Exx, \Box\forall y(y \prec^{\Diamond} xx), (\mathrm{S}^-), \forall v\phi_1 \vdash \forall v(v \prec^{\Diamond} xx \to \Box\phi_1^{\Diamond}) \qquad \mathsf{FPFO}$$

$$Exx, \Box\forall y(y \prec^{\Diamond} xx), (\mathrm{S}^-), \forall v\phi_1 \vdash \Box\forall v(v \prec^{\Diamond} xx \to \phi_1^{\Diamond}) \qquad \text{RBF}$$

$$Exx, \Box\forall y(y \prec^{\Diamond} xx), (\mathrm{S}^-), \forall v\phi_1 \vdash \Box\forall v(v \prec^{\Diamond} xx) \to \Box\forall v\phi_1^{\Diamond} \qquad \mathsf{FPFO}\text{, K}$$

$$Exx, \Box\forall y(y \prec^{\Diamond} xx), (\mathrm{S}^-), \forall v\phi_1 \vdash \Box\forall v\phi_1^{\Diamond} \qquad \mathsf{FPFO}\text{, K}$$

(iv) $\phi = \forall vv\phi_1$; $\phi^{\Diamond} = \Box\forall vv\phi_1^{\Diamond}$. Replacing RBF with RBF_p in the argument for (iii), we obtain

$$Exx, \Box\forall yy(yy \prec^{\Diamond} xx), (\mathrm{S}^-), \forall vv\phi_1 \vdash \Box\forall vv\phi_1^{\Diamond} \qquad \mathsf{FPFO}\text{, K}$$

The result follows from (13) $\square$

Proof of (F**).* The proof adapts the proof of (F). Replacing Φ with $\Phi^\star$ trivialises the atomic case, (i) The other cases (ii)–(iv) proceed as before dropping (S^-) as an assumption. $\square$

5 Modal Set Theory[1]

Relativist accounts of what goes wrong with the set-theoretic paradoxes are seldom accompanied by an account of what goes right with Zermelo-Fraenkel set theory. Indeed sometimes the means to cure set theory of the paradoxes threatens to seriously disable the patient. Predicative solutions to the paradoxes propose that we reign in plural or higher-order Comprehension axioms in order to retain Collapse.[2]

Collapse. $\forall xx \exists y \forall z (z \in y \leftrightarrow z \prec xx)$

Predicative Comprehension. $\exists xx \forall z (z \prec xx \leftrightarrow \phi)$

In the plural case, Predicative Comprehension is the schema that adds the further side condition that ϕ lacks any plural quantifiers (in addition to the existing side constraint that ϕ lack xx). In languages which contain $\in$ as an atomic predicate, however, this does not suffice to restore consistency. The pathological instance of Plural Comprehension where $z \notin z$ replaces ϕ is not yet ruled out. To avoid this problem, the friend of predicativism must impose the further side condition that ϕ lack $\in$. (Note that in languages containing a function symbol $\{\cdot\}$ that yields a singular term $\{\tau\tau\}$ when supplied with a plural term $\tau\tau$ and $x \in y$ is a defined expression:

$$x \in y =_{\text{df}} \exists yy (y = \{yy\} \wedge x \prec yy)$$

the pathological instance is already ruled out by the ban on plural variables.)

[1]A later version of this chapter is forthcoming as 'The iterative conception of set: a (bi-)modal axiomatisation' in the *Journal of Philosophical Logic*. The final publication is available at www.springerlink.com.

[2]Dummett (1991, ch. 17) blames impredicativity for the paradoxes in Frege's logic.

This retreat to weaker comprehension axioms restores consistency.[3] But, as John Burgess (2005, ch. 2) has demonstrated in a similar setting, the cost to logical strength is severe. The resulting theory is far too weak to recover Peano Arithmetic let alone anything approaching Zermelo-Fraenkel set theory.

To strive for further strength the predicativist may move to a ramified system. First we replace the unsorted plural variables $xx, yy, zz, \ldots$ with plural variables of countably many different sorts:

$$\begin{gathered} xx^0, yy^0, zz^0, \ldots \\ xx^1, yy^1, zz^1, \ldots \\ xx^2, yy^2, zz^2, \ldots \\ \vdots \end{gathered}$$

Similarly the element-set predicate $\in$ is replaced by countably many indexed predicates $\in^0, \in^1, \ldots$. Collapse and Predicative Comprehension are then replaced with the following principles.

Ramified Collapse. $\forall xx^i \exists y \forall z (z \in^i y \leftrightarrow z \prec xx^i)$

Ramified Comprehension. $\exists xx^i \forall z (z \prec xx^i \leftrightarrow \phi_i)$

Ramified Comprehension is subject to the side constraints that ϕ_i (i) only contain free plural variables xx^j for $j \leqslant i$; (ii) only contain bound plural variables xx^k for $k < i$; and (iii) only contain $\in^k$ for $k < i$. (Again, in languages where the function symbol $\{\cdot\}$ is taken as primitive in place of the element-set predicates, which are reintroduced by the definitions:

$$x \in^k y =_{\mathrm{df}} \exists yy^k (y = \{yy^k\} \wedge x \prec yy^k),$$

the third condition (iii) is automatically enforced by adherence to (ii).)

[3] Heck (1996) shows that predicative second-order logic together with Frege's Basic Law V is consistent. The predicative plural system in the text is readily interpreted in this system via the definition of $\in$ in the text and is, therefore, also consistent.

However, as Burgess shows, ramification does little to alleviate the severe loss of strength. The ramified set theory still falls far short of Peano Arithmetic.[4] This is an exorbitant cost to pay for retaining a maximally liberal attitude to which objects are able to form a set. The business of this chapter will be to show that the relativist need not pay it. When given a modal articulation, the maximally liberal attitude gives rise to a very natural motivation for standard set theory, via a modal formulation of the iterative conception of set.

1 The Iterative Conception.

Recall the iterative conception of set, brought to prominence amongst the philosophical community by George Boolos (1971). Here is the version for pure sets. (In this chapter 'set' means pure set.)[5] At the very first stage—we start counting at stage 0—there are no sets, and none are formed. At each subsequent stage, all sets of sets formed at earlier stages will be formed. At the second stage, stage 1, $\varnothing$ will be formed; at stage 2, $\varnothing$ and $\{\varnothing\}$ will be formed; at stage 3, $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$ and $\{\varnothing, \{\varnothing\}\}$ will be formed; similarly for stages 4, 5, ... Immediately after these stages will come the first limit stage, stage ω, when all sets of sets formed at finite stages will be formed. At this stage, all the hereditarily finite sets will be formed. At the next stage, stage $\omega + 1$, the first sets of infinite rank will be formed. And so it goes on, stretching as high as the ordinals are long.

Beyond the structure of the stages, and the principle of Extensionality—that sets with the same elements are identical—the iterative conception thus amounts to two key principles. The first articulates the priority of a set's elements to the set itself; the

[4] The systems Burgess (2005, ch. 2) focuses on are higher-order rather than plural and contain Basic Law V rather than Collapse; but since these systems clearly interpret the ones in the text (via the displayed definitions of $\in$ and $\in^k$) the weakness results immediately carry over. Burgess shows that the predicative theory is unable to interpret Robinson Arithmetic together with induction restricted only to formulas with bounded quantifiers and axioms defining exponentiation and the ramified theory is unable to interpret the result of adding definitions of two further arithmetical operations to this theory: 'superexponentiation' and 'supersuperexponentiation'.

[5] It would be straightforward to adapt the theory developed in this chapter to recover ZFCU but we here follow the common set theoretic practice of ignoring urelements.

second concerns the plenitude of sets.

Priority. Any set ever formed is formed from some zero or more sets that were formed at an earlier stage.

Plenitude. Any zero or more sets formed at any stage will form a set at every later stage.[6]

Boolos's view is that neither the tense in this informal sketch of the iterative conception, nor the metaphor of set 'formation', should be taken at all seriously. The informal sketch is just a colourful story or 'narrative convention'[7] that ultimately gives way to a formal axiomatisation of the iterative conception, in the form of a 'stage theory'.

Boolos axiomatises stage theory in a nonmodal, two-sorted language, with quantification over sets $(x, y, \ldots)$ and stages $(s, t, \ldots)$. The tense is eliminated in favour of quantification over stages, and the formation metaphor dispensed with by taking the predicates *s is earlier than t* and *x is formed at s* as primitive, governed by certain stage-theoretic axioms, which formalise the informal sketch of the conception.[8]

Such a theory is not without its merits. Indeed, it represents one of the best attempts to date to find a natural motivation for a good deal of standard set theory, permitting us to derive many of the ZF axioms from a theory that appears not at all ad hoc.[9] But the trouble with formalising stage theory in a non-modal language is that we lose the ability to sustain the compelling thesis that all that is required of the world for some sets to be able to form a set is that there be those sets. The maximally liberal attitude to set formation *is* upheld by the tensed sketch of the iterative conception. No matter what stage has been reached, any sets formed then *can* form a set—indeed, *will* form a set at the very next stage. But, no matter how far we have proceeded up the hierarchy, the sets formed so far are *all* the sets there are: (otherwise) unrestricted

[6] As ever plural quantifiers $\forall vv$ are to be read 'for any zero or more things...'.

[7] Boolos (1989), p. 90, attributes the thought that the iterative conception is a narrative convention, describing of the set-theoretic hierarchy, according to one salient aspect of its structure—rank by rank—to Dan Leary.

[8] Boolos gives two axiomatisations of stage theory. See Boolos (1971), pp. 20–22, and Boolos (1989), p. 91.

[9] Just which axioms is controversial; in particular, it has been disputed—notably by Boolos (1989)—whether the iterative conception motivates the axioms of Extensionality and Replacement. See Alexander Paseau (2007) for discussion.

quantification over sets ranges only over the ones formed so far.[10] So understood, the iterative conception sustains what we shall call the *Liberalness thesis*: Any sets *can* form a set.

Once we eliminate the tense, however, this thesis collapses into the principle that any sets whatsoever *already* form a set, which leads to inconsistency in the familiar way via Russell's paradox. Accordingly, as far as we can make sense of the question of which sets are able to form a set in a tenseless, nonmodal theory, we must give a less than maximally liberal response; for instance, the sets that lack themselves as elements do not form a set. Such a response raises a difficult question: *What is it about the world that allows some sets to form a set, whilst prohibiting others from doing the same?* As we saw in the previous chapter no satisfactory response has been forthcoming from the absolutist.

The relativist, however, is able to avoid such difficulties by upholding Liberalness: contending that any things *can* form a set—in the relevant sense. But this also provides a means to take the tense in the informal outline of the iterative conception more seriously than is customary. Rather than eliminating the tense in favour of quantification over stages, it may instead be replaced with suitable inter-perspective modal operators governed by a tense-*like* logic. This chapter will show that such a stage theory allows for a natural development of standard set theory.

A similar approach was pioneered by Charles Parsons (1977, 1983) and more recently further developed by Øystein Linnebo (2010, n.d.). Both theories enrich the language of set theory with a single modal operator (governed by an extension of $\mathsf{S4.2}$)[11]

---

[10] There are no doubt many—four dimensionalists, for instance, or friends of the Barcan formula for tense operators—who would dispute this assumption about the range of unrestricted quantifiers in tensed language. Accordingly, let us hedge, and say that the tensed sketch sustains the thesis *if quantifiers have this property.* That this sketch supports this thesis is ultimately unimportant; what matters is that the stage theory it prompts upholds it.

[11] Recall that, for necessity-operator $\Box$ and corresponding possibility-operator $\Diamond$, $\mathsf{S4}$ adds the axioms (T) $\Box\phi \to \phi$ and (4) $\Box\phi \to \Box\Box\phi$ to a normal system, corresponding to the accessibility relation's being reflexive and transitive; $\mathsf{S4.2}$ additionally adds (G) $\Diamond\Box\phi \to \Box\Diamond\phi$, corresponding to its being directed:

$$\forall x_1 \forall x_2 \forall x_3 (Rx_1x_2 \wedge Rx_1x_3 \to \exists x_4 (Rx_2x_4 \wedge Rx_3x_4)).$$

and use this language to develop what may be seen as a modal stage theory that sustains Liberalness; before showing how (a modal analogue of) standard set theory may be recovered in it.

While these theories offer a great improvement over nonmodal stage theories, such a unimodal language lacks the expressive resources to fully articulate the iterative conception—lacking, in particular, the means to characterise the well-foundedness of the earlier-later relation on stages—with the result that certain set-theoretic principles—most notably the axiom of foundation—must simply be added to the stage theory in order to secure their recovery. To overcome this deficiency, I will present a bimodal formulation of stage theory. The addition of a further modal operator allows for a very natural axiomatisation of the iterative conception, allowing us to derive as theorems what would otherwise need to be taken as basic.

Bimodal stage theory is formulated in a language that adds two modal operators to the language of set theory: the 'forwards-looking' and 'backwards-looking' necessity-operators $\Box_>$ and $\Box_<$. Pending further elaboration on their intended interpretation, these operators may be provisionally read as follows:

$\Box_>\psi$	'it will be the case at every later stage that ψ.'
$\Box_<\psi$	'it was the case at every earlier stage that ψ.'

Other modalities may then be introduced via abbreviations in the metalanguage. An 'absolute' necessity-operator $\Box\psi$ ('it is always the case at every stage that ψ') is defined as $\Box_<\psi \wedge \psi \wedge \Box_>\psi$; a 'weakly forward-looking' necessity-operator (which is the analogue of the single modal operator employed by Parsons and Linnebo) is defined such that $\Box_\geqslant\psi =_{\mathrm{df}} \Box_>\psi \wedge \psi$, similarly for the 'weakly backwards-looking' necessity-operator: $\Box_\leqslant\psi =_{\mathrm{df}} \Box_<\psi \wedge \psi$. The corresponding possibility-operators are then introduced in the usual way, for instance: $\Diamond\psi =_{\mathrm{df}} \neg\Box\neg\psi$.

Our target is first-order Zermelo Fraenkel set theory, ZF. The modal stage theory formulated in this language is presented in section 4, together with an outline of how the modal analogue of a natural extension of Zermelo set theory less the axiom of Infinity may be recovered from it, and how it may be elaborated to recover the modal analogue of ZF. First, however, two preliminaries are in order. Section 2 motivates and presents the logic of the modality in question and elucidates its intended interpretation. Section 3 assesses the relationship between the modal analogues of theories and their more familiar nonmodal counterparts and presents a few useful results.

2 Mathematical Modality.

The logic of stage theory places general constraints on how sets may be formed in successive stages in accordance with the iterative conception,[12] without itself postulating the existence of any particular set. Rather than attempting to motivate the logic of stage theory directly, however, we shall start with a Kripke-style model theory.[13]

2.1

The basic set up gives some suggestive labels to some familiar machinery. Let $\mathcal{L}^{\Diamond}$ be the bimodal language that adds $\Box_{>}$ and $\Box_{<}$ to a first-order language $\mathcal{L}$, with identity. A *skeleton* $\mathcal{S}$ is a pair $\langle S, < \rangle$. The 'hierarchy of stages' S is a non-empty set and the 'earlier-later relation' $<$ is a binary relation on S. For each stage s in S, a *frame* $\mathcal{F}$ based on $\mathcal{S}$ adds a set M_s to $\mathcal{S}$, the 'universe of sets formed so far' or 'stage-universe', at least one of which is non-empty. For each n-ary predicate P in $\mathcal{L}^{\Diamond}$ and each s in S, a *model* $\mathcal{M}$ based on $\mathcal{F}$ adds to $\mathcal{F}$ an n-ary relation $|P|_s$ on M_s as the 'stage-extension' of P. The stage-extension of the identity predicate remains fixed, with $|{=}|_s = \{\langle a, a\rangle \mid a \in M_s\}$. An *assignment* σ over the model $\mathcal{M}$ is a function that maps each variable v to an element $\sigma(v)$ of any stage-universe M_s. Relative to a model, that a formula ψ is 'true at stage s

[12]The logic of stage theory need not be conceived as specific to *set* formation. We might also conceive of other objects, for instances cardinals and ordinals (construed *sui generis*), as formed in the series of stages governed by this logic.
[13]See John P. Burgess (2002) and Yde Venema (2001) for recent overviews of the model theory and proof theory of such logics.

under assignment σ' (in symbols: $|\psi|_s{}^\sigma = T$) is defined recursively in the usual way.[14]

$$|Rv_1, \ldots, v_k|_s{}^\sigma = T \text{ if and only if } \langle\sigma(v_1), \ldots, \sigma(v_k)\rangle \in |R|_s;$$
$$|\neg\psi|_s{}^\sigma = T \text{ if and only if } |\psi|_s{}^\sigma \neq T;$$
$$|\psi_1 \to \psi_2|_s{}^\sigma = T \text{ if and only if } |\psi_1|_s{}^\sigma \neq T \text{ or } |\psi_2|_s{}^\sigma = T;$$
$$|\forall v\psi|_s{}^\sigma = T \text{ if and only if, for every } a \in M_s, |\psi|_s{}^{\sigma[v/a]} = T;$$
$$|\Box_{>}\psi|_s{}^\sigma = T \text{ if and only if, for every } t > s, |\psi|_t{}^\sigma = T;$$
$$|\Box_{<}\psi|_s{}^\sigma = T \text{ if and only if, for every } t < s, |\psi|_t{}^\sigma = T.$$

(Here, the 'later-earlier' relation, $s > t =_{\text{df}} t < s$.)

Notice that we take an actualist attitude to quantifiers and predicates. At any stage $\forall v$ ranges only over the sets formed then; and since $|P|_s$ is a relation *on* M_s, $Pv_1, \ldots, v_k$ is true then only if $v_1, \ldots, v_k$ denote sets formed then.

We also obtain the expected truth-conditions for the defined necessity-operators.

$$|\Box_{\geqslant}\psi|_s{}^\sigma = T \text{ if and only if, for every } t \geqslant s, |\psi|_t{}^\sigma = T;$$
$$|\Box_{\leqslant}\psi|_s{}^\sigma = T \text{ if and only if, for every } t \leqslant s, |\psi|_t{}^\sigma = T;$$
$$|\Box\psi|_s{}^\sigma = T \text{ if and only if, for every } t \lesseqgtr s, |\psi|_t{}^\sigma = T.$$

(Here, $s \leqslant t =_{\text{df}} s < t \vee s = t$; $s \geqslant t =_{\text{df}} s > t \vee s = t$; and $s \lesseqgtr t =_{\text{df}} s \leqslant t \vee t \leqslant s$.)

As usual a formula is said to be *valid in a model* $\mathcal{M}$ if it is true at every stage of $\mathcal{M}$ under any assignment over it; *valid in a frame* $\mathcal{F}$ if valid in every model based on $\mathcal{F}$ and *valid in a skeleton* $\mathcal{S}$ if valid in every frame based on $\mathcal{S}$

With this model theory in place, the iterative conception gives rise to eight further constraints that models faithful to it should satisfy. The first six concern the ordering of the stages. As is implicit in their ordinal indexing, the iterative conception has it that the stages are well-ordered by the earlier-later relation, and that there is no last stage. Accordingly, a faithful skeleton need be:

[14]Notation: the assignment $\sigma[v/a]$ assigns a to v and otherwise agrees with σ.

(i) transitive: $(\forall s \in S)(\forall t \in S)(\forall u \in S)\big(s < t \wedge t < u \rightarrow s < u\big)$;

(ii) well-founded: $(\forall U \subseteq S)\big(U \neq \varnothing \rightarrow (\exists u \in U)(\forall s \in U)(\neg s < u)\big)$
(and therefore also irreflexive);

(iii) connected: $(\forall s \in S)(\forall t \in S)(s \leqslant t \vee t \leqslant s)$ and therefore, in particular:

  (iv) without forward branching: $(\forall s \in S)(\forall t_1 > s)(\forall t_2 > s)(t_1 \leqslant t_2 \vee t_2 \leqslant t_1)$;

  (v) without backward branching: $(\forall s \in S)(\forall r_1 < s)(\forall r_2 < s)(r_1 \leqslant r_2 \vee r_2 \leqslant r_1)$;

(vi) serial: $(\forall s \in S)(\exists t \in S)(t > s)$.

The next constraint concerns the stage-universes. According to the iterative conception, once a set is formed it is never destroyed; it remains formed at every subsequent stage. Later stage-universes are never less inclusive than earlier ones. Consequently, a faithful frame is also:

(vii) monotonic: $(\forall s \in S)(\forall t \in S)(s < t \rightarrow M_s \subseteq M_t)$.

Lastly the iterative conception prompts a constraint on the relation between different stage-extensions of $\in$. On this conception, when new sets are formed, the element-set relations between the sets formed already are left unchanged: elements of antecedently formed sets remain elements and non-elements of such sets remain non-elements. Consequently different stage-extensions of $\in$ agree on their common domains. More generally, a faithful model is:

(viii) stable: $(\forall s \in S)(\forall t_1 \leqslant s)(\forall t_2 \leqslant s)(|P|_{t_1} \cap M_{t_2}^n = |P|_{t_2} \cap M_{t_1}^n)$.

While the truth-conditions of $\mathcal{L}^{\Diamond}$-formulas in models that satisfy these constraints no doubt give us some sort of insight into their intended interpretations, it bears some emphasis that the model theory should not be taken *too* seriously. It cannot be maintained that the modality is semantically reducible to quantification over stages without modal stage theory collapsing into nonmodal stage theory, and thereby losing the ability to sustain Liberalness. Consequently, we shall develop modal stage theory proof-theoretically and restrict our use of Kripke models to motivating its logical axioms, and other instrumental purposes.

Table 5.1: LST axioms and rules, first two groups

FUS	$\vdash \forall v\psi \to (Eu \to \psi[u/v])$
FUG	$\Delta, Eu \vdash \psi_1 \to \psi_2[u/v]$ only if $\Delta \vdash \psi_1 \to \forall v\psi_2$
Refl	$\vdash \Phi(v) \to v = v$
Sub	$\vdash u = v \to (\psi \to \psi(u/v))$
$\mathrm{K}_<$	$\vdash \Box_<(\psi_1 \to \psi_2) \to (\Box_<\psi_1 \to \Box_<\psi_2)$
$\mathrm{K}_>$	$\vdash \Box_>(\psi_1 \to \psi_2) \to (\Box_>\psi_1 \to \Box_>\psi_2)$
$\mathrm{Nec}_<$	$\vdash \psi$ only if $\vdash \Box_>\psi$
$\mathrm{Nec}_>$	$\vdash \psi$ only if $\vdash \Box_<\psi$
$\mathrm{CV}_<$	$\vdash \psi \to \Box_<\Diamond_>\psi$
$\mathrm{CV}_>$	$\vdash \psi \to \Box_>\Diamond_<\psi$

FUG is subject to the side constraint that u does not occur free in $\psi_1 \to \forall v\psi_2$.

2.2

The logic of modal stage theory is a bimodal first-order logic, which we call LST. In addition to a standard axiomatisation of nonmodal propositional logic, the axioms and rules of LST fall into three natural groups. (See tables 5.1 and 5.2.)

The first four axioms and rules govern quantification and identity on their actualist reading.[15] Since stage-universes vary, and quantifiers range only over the sets formed so far, we use a free logic as in chapter 4. The first four axioms, are the first-order axioms of the free plural logic outlined there, FPFO. (See table 4.1 on page 112.) As before, $Ev =_{\text{df}} v = v$.

The second group of axioms and rules are those of basic tense logic, which add to the axioms characteristic of a normal system for forwards- and backwards-necessity, two more governing their interaction. It is readily verified that the axioms in the first two groups are valid over all skeletons.

It is the final eight axioms that articulate the distinctive structure of the hierarchy of stages according to the iterative conception. These axioms respectively characterise the transitivity, well-foundedness,[16] connectedness, lack of forwards and backwards

[15] Notation: when v is free for u, $\psi[u/v]$ is the result of replacing every free occurrence of u in ψ with v, $\psi(u/v)$ the result of replacing any (or no) such occurrences. As far as possible we adopt the labelling of Bull and Segerberg (2001), with the obvious convention on subscripts.

[16] In the context of provability logic, Boolos observes the correlation between Löb's principle: $\Box(\Box\psi \to \psi) \to \Box\psi$ and the transitivity and converse well-foundedness of the accessibility relation. Boolos (1995, pp. 75–6). Thanks to Gabriel Uzquiano for drawing my attention to the possibility of a backwards-looking version of this principle as a means to express well-foundedness.

Table 5.2: LST axioms, third group

4	$\vdash \Box_{<}\psi \to \Box_{<}\Box_{<}\psi$	(I)
WF	$\vdash \Diamond\psi \to \Diamond(\psi \wedge \Box_{<}\neg\psi)$	(II)
E_1	$\vdash \Diamond Ev$	(III)
$\mathrm{H}_{>}$	$\vdash \Diamond_{>}\psi_1 \wedge \Diamond_{>}\psi_2 \to \big(\Diamond_{>}(\psi_1 \wedge \psi_2) \vee \Diamond_{>}(\psi_1 \wedge \Diamond_{>}\psi_2) \vee \Diamond_{>}(\psi_2 \wedge \Diamond_{>}\psi_1)\big)$	(IV)
$\mathrm{H}_{<}$	$\vdash \Diamond_{<}\psi_1 \wedge \Diamond_{<}\psi_2 \to \big(\Diamond_{<}(\psi_1 \wedge \psi_2) \vee \Diamond_{<}(\psi_1 \wedge \Diamond_{<}\psi_2) \vee \Diamond_{<}(\psi_2 \wedge \Diamond_{<}\psi_1)\big)$	(V)
$\mathrm{D}_{>}$	$\vdash \Box_{>}\psi \to \Diamond_{>}\psi$	(VI)
$\mathrm{CBF}_{>}$	$\vdash \Box_{>}\forall v\psi \to \forall v\Box_{>}\psi$	(VII)
STA	$\vdash \Box(E_\Phi \to \Phi) \vee \Box(E_\Phi \to \neg\Phi)$	(VIII)

branching and seriality of the earlier-later relation, the monotonicity of stage-universes and the stability of stage-extensions in the following way.[17]

Proposition 1.

(a) (I) is valid on a skeleton iff it satisfies (i)

(b) The same is true of (II)–(VI) and (ii)-(vi)

(c) (VII) is valid on a frame iff it satisfies (vii)

(d) (VIII) is valid in a model iff it satisfies (viii) $\Box$

Derivability is characterised in the usual way;[18] and soundness is readily established on the basis of proposition 1: $\vdash \psi$ only if ψ is valid in all models satisfying (i)–(viii). The bimodal system LST permits us to derive the K principle and the rule of necessitation for each defined modal operator, $\Box$, $\Box_{\leqslant}$ and $\Box_{\geqslant}$. Furthermore, we may derive the following unimodal principles for these operators.

Proposition 2 (Derived unimodal principles)**.**

(a) $\Box$ conforms to the axioms characteristic of $\mathsf{S5}$: T: $\Box\psi \to \psi$ (corresponding to the reflexivity of $\leqslant$), *B*: $\psi \to \Box\Diamond\psi$ (symmetry) and 4: $\Box\psi \to \Box\Box\psi$ (transitivity). As usual we may derive *E*: $\Diamond\psi \to \Box\Diamond\psi$ (corresponding to Euclideanness).[19]

[17] Several of the propositional cases are covered by Goldblatt (1992) thms. 1.12–3.

[18] A formula ψ is *derivable* from Δ in LST (in symbols: $\Delta \vdash_{\mathsf{LST}} \psi$) if ψ is the last member of a finite sequence of formulas each of which is an axiom, an element of Δ, or the result of applying a rule to earlier members in the sequence; with UG subject to the constraint that save for Eu earlier members of the sequence and $\psi_1 \to \forall v\psi_2$ should not contain u free, and $\mathrm{Nec}_{<}$ and $\mathrm{Nec}_{>}$ subject to the constraint that no elements of Δ may occur earlier in the sequence.

[19] Recall that a relation R is said to be *Euclidean* if $\forall x\forall y\forall z(Rxy \wedge Rxz \to Ryz)$.

(b) $\Box_{\leqslant}$ conforms to the axioms characteristic of $\mathsf{S4.3}$: $\mathrm{T}_{\leqslant}$, $4_{\leqslant}$ (reflexivity and transitivity of $\leqslant$) and $\mathrm{H}_{\leqslant}$: $\Diamond_{\leqslant}\psi_1 \wedge \Diamond_{\leqslant}\psi_2 \rightarrow \big(\Diamond_{\leqslant}(\psi_1 \wedge \psi_2) \vee \Diamond_{\leqslant}(\psi_1 \wedge \Diamond_{\leqslant}\psi_2) \vee \Diamond_{\leqslant}(\psi_2 \wedge \Diamond_{\leqslant}\psi_1)\big)$ (lack of backwards branching). (We apply the obvious convention for naming principles: for example, $\mathrm{T}_{\leqslant}$ is the principle: $\Box_{\leqslant}\psi \rightarrow \psi$.) This operator also conforms to the Barcan formula, $\mathrm{BF}_{\leqslant}$: $\forall v \Box_{\leqslant}\psi \rightarrow \Box_{\leqslant}\forall v \psi$ (corresponding to earlier stage-universes' being no more inclusive than later ones).

(c) $\Box_{\geqslant}$ likewise conforms to the principles characteristic of $\mathsf{S4.3}$: $\mathrm{T}_{\geqslant}$, $4_{\geqslant}$ (reflexivity and transitivity of $\geqslant$) and $\mathrm{H}_{\geqslant}$ (lack of forwards branching). This operator also conforms to $\mathrm{CBF}_{\geqslant}$: $\Box_{\geqslant}\forall v \psi \rightarrow \forall v \Box_{\geqslant}\psi$ (corresponding to later stage-universes' being no less inclusive than earlier ones). $\Box$

Finally, we record some familiar derived rules that will be used freely and often tacitly in our development of modal stage theory below.[20]

Proposition 3 (Derived rules)**.** Let ψ_1 and ψ_2 be $\mathcal{L}^{\Diamond}$-formulas and Δ a set of these.

(a) Deduction theorem. If $\Delta, \psi_1 \vdash \psi_2$, then $\Delta \vdash \psi_1 \rightarrow \psi_2$.

(b) Substitution theorem. Let $\theta(\psi/\Phi)$ be the result of uniformly substituting ψ for the atomic formula Φ in formula θ. Suppose that, for each δ in Δ, $\Delta \vdash O\delta$ for each modal operator O in θ and δ does not have free any variable bound in θ.

$$\text{If } \Delta \vdash \psi_1 \leftrightarrow \psi_2, \text{ then } \Delta \vdash \theta(\psi_1/\Phi) \leftrightarrow \theta(\psi_2/\Phi).$$

(c) Becker's rule. Let each O_j be either a modal operator of the form L_j or $\neg L_j \neg$ for one of the five (simple or defined) necessity-operators or a quantifier of the form $\forall v$ or $\exists v$, for $j = 1, \ldots n$. Suppose that, for each δ in Δ, for each modal operator O_j, $\Delta \vdash L_j\delta$, and for each quantifier O_j, δ does not have free any variable v that occurs in O_j, for $j = 1, \ldots, n$.

$$\text{If } \Delta \vdash \psi_1 \rightarrow \psi_2, \text{ then } \Delta \vdash O_1 \ldots O_n\psi_1 \rightarrow O_1 \ldots O_n\psi_2 \qquad \Box$$

[20]These are proved in the usual way. For the last two, see, for instance, Burgess (2002, p. 8).

2.3

The time has come to say something further about the intended interpretation of the modal operators. Taking the tense in the iterative conception seriously may provoke two familiar objections.

(a) *Too few moments.* The stages are as numerous as the ordinals. They exceed any infinite cardinality $\aleph_\alpha$ in number. But there is no reason to think time (or spacetime) has a structure even approaching this complexity. Cardinality consideration immediately rule out our embedding the hierarchy of stages within the succession of moments (or the manifold of spacetime).[21]

(b) *Wrong modal profile.* New sets are formed at each stage. And thus some sets formed at the present stage did not exist at earlier stages. But this is to get the modal profile of sets wrong: (pure) sets exist necessarily if possibly and eternally if ever.[22] Modal considerations immediately rule out our bringing sets into existence.

Both of these constitute powerful objections against a *face value* reading of the iterative conception, where $\Box_<$ and $\Box_>$ are read as tense operators and talk of 'formation' is construed literally.

However, this is *not* the relativist's proposal. Rather the modal operators are construed as inter-perspective modal operators like those presented in chapter 3 (section 4). Stages may be identified with certain perspectives: namely, those whose domains comprise the elements of a cumulative rank V_α. Call such perspectives ranklike. The two modal operators may then be given a less metaphorical gloss as follows.

$\Box_>\psi$ 'ψ holds at every ranklike perspective with a wider domain.'

$\Box_<\psi$ 'ψ holds at every ranklike perspective with a narrower domain.'

[21]Parsons (1977) raises worries of this sort against Hao Wang's account of the iterative conception in Wang (1974), ch. 6

[22]See, for instance, Kit Fine (1981).

(As ever the pluralist will accept this gloss only if the right hand-side is taken to have implicit modal content. See chapter 3, section 4.1.)[23]

Relativists of different stripes will disagree over the interpretation of the modal operators according to whether shifts in perspectives are due to shifts in context, interpretation or third-parameter. But what is common to each of these views is that the shifts are non-circumstantial. And such non-circumstantial interpretations of the modal operators avoid the problems for the face-value, temporal interpretation.

Against (a), unlike spacetime, there seems no reason to think that the space of possible interpretations, or the space of possible perspectives is impoverished in comparison to the ordinals. And at least in the case of the interpretational relativist, there is every reason to think that there are as many possible interpretations as ordinals. For it is possible for a constant symbol to denote any ordinal and interpretations that differ over the denotation of a constant symbol are *ipso facto* distinct.

Against (b), the perspective parameter p is orthogonal to the world parameter w. In order to literally bring the empty set into existence, for example, we would have to bring about a change in the world, from w_1 to w_2, say, such that the sentence

$$\exists y(y = \varnothing)$$

is false in w_1 and true in w_2 (relative to a fixed perspective). But none of the relativist accounts are committed to any such thing. Rather the 'formation' of the empty sets consists in a shift of perspective—whether this is a shift of context, or interpretation, or third parameter—from p_1 to p_2, such that this sentence is false under the first interpretation and true under the second one (leaving the circumstances w unchanged). The same is true when we replace the term $\varnothing$ with a term denoting any other set. The shifts in such existential sentences' truth-value under p-shifts from one ranklike perspective to another (relative to a fixed possible world w) posited on these accounts are compatible with the invariance of their truth-value under w-shifts called for by the pre-theoretic

[23]Under this interpretation, the defined modal operator $\Box$ of this chapter expresses a more restricted modality than its primitive homonym $\Box$ of earlier chapters. The latter is not restricted to ranklike perspectives. The logics of the two modal operators, however, are the same.

judgements about the modal profile of sets behind objection (b).

The non-circumstantial interpretations open to relativists avoid the problems facing a face value interpretation of the iterative conception. Nonetheless in view of the logical similarities between tense logic and LST, I shall persist in employing tensed language and the *metaphor* of 'formation' as a heuristic tool in informal explanations. (Of course, all formal arguments are carried out within LST and in no way rely on this gloss.)

3 Modal Mathematics.

Set theory is conspicuously not standardly formulated in a modal language. In order, therefore, for modal stage theory to motivate standard set theory we need to explain what the relationship is between the modal theory and its more familiar non-modal counterpart. Here we encounter a third *prima facie* concern.

(c) *Insufficient generality.* On the modal account, no matter how far we ascend up the hierarchy, our quantifiers only range over the sets formed so far. This leaves us unable to express set-theoretic claims in their full, intended generality. For example, when a set theorist attempts to express the principle of foundation by uttering 'Every non-empty set has an $\in$-minimal element' she succeeds only in saying that every non-empty set *formed at the current stage* has an $\in$-minimal element, leaving open the possibility of non-well-founded sets being formed at other stages.

As in chapter 3, pluralist and thorough-going relativists will respond in different ways to this objection.[24] The pluralist contends that the fact that our quantifiers only ever range over the sets formed at the current stage does not prevent our generalising about absolutely every set ever formed. This is achieved by embedding set-theoretic quantifiers within inter-perspective modal operators. In terms of Kripke models, while $\forall v\psi$ says only that every member of the current stage satisfies ψ, $\Box\forall v\psi$ say that every member of every stage satisfies ψ. Likewise, while the extension of $u \in v$ may vary from stage to

[24]This is analogous to the sort of problems raised for the relativist about quantification discussed in chapter 3 (section 3).

stage, gaining new element-set pairs as new sets are formed, $\Diamond(u \in v)$ is *always* satisfied by every element-set pair. As in chapter 4 (section 2.3) we define $\phi^{\Diamond}$, the *modalisation* of an $\mathcal{L}_{\in}$-formula ϕ, as follows.[25] For atomic formulas: $(u = v)^{\Diamond} =_{\mathrm{df}} \Diamond(u = v)$ and $(u \in v)^{\Diamond} =_{\mathrm{df}} \Diamond(u \in v)$. To tidy the notation, the modalisations of atomic formulas are written as $x \in^{\Diamond} y$, $x =^{\Diamond} y$, and so on. For complex formulas: $(\neg\phi)^{\Diamond} = \neg\phi^{\Diamond}$; $(\phi_1 \to \phi_2)^{\Diamond} = \phi_1^{\Diamond} \to \phi_2^{\Diamond}$; $(\forall v\phi)^{\Diamond} = \Box\forall v\phi^{\Diamond}$.

The pluralist may then claim that the intended generality of set theory—as a theory concerning every set ever formed—is attained by treating the utterances of set theorists as having implicit modal content.[26] The intended generality of a statement made in the nonmodal language $\mathcal{L}_{\in}$ is most perspicuously expressed in $\mathcal{L}_{\in}^{\Diamond}$ by its modalisation. For instance, the intended generality of the principle of foundation is captured by the following modalised formula, which rules out non-well-founded sets ever being formed:

$$\Box\forall x\big(\Diamond\exists z(z \in^{\Diamond} x) \to \Diamond(\exists z \in^{\Diamond} x)\Box(\forall w \in^{\Diamond} z)(w \notin^{\Diamond} x)\big).$$

The thorough-going relativist can accept this sort of line up to a point. On her view, modalised quantifiers no more achieve absolute generality than quantifiers do. Just as quantifiers can always be more inclusively interpreted to range over new objects, so too can inter-perspective modal operators be more liberally interpreted to encompass new perspectives. However, she also denies that the *intended* generality of set theory is absolute generality. Even when we are engaged in pure set theory, it is enough for quantifiers or modalised quantifier to encompass enough sets for the purpose at hand. Consequently, the thorough-going relativist may also accept that, when suitable interpreted, modalised quantifiers succeed in capturing the intended generality of set theory, even if these are not absolutely general. (She also contends, however, that the same generality could be achieved without modal operators, with quantifiers interpreted relative to a perspective

[25] Linnebo gives a similar definition of modalisation according to which the necessity operator he uses (the equivalent of $\Box_{\geqslant}$) is prefixed to quantifiers as here, but no modal operators are prefixed to atomic formulas. (See Linnebo (2010,n.d.).) This would be problematic in the current setting since the truth-values of atomic formulas are not stage-invariant. (Such formulas are however ‘stable’, in Linnebo’s sense, see n. 28.)

[26] Linnebo (2010) makes the corresponding suggestion for his version of modalisation. See also Fine (2006), pp. 41–3.

outside the range of the modal operators).

But on either account, it will only be plausible to interpret the utterances of mathematicians, or a mathematical theory given in $\mathcal{L}_\in$, as having such modal content if it is consonant with mathematical practice to do so. The next two subsections argue that this is the case.[27]

3.1

One feature we should expect from modalised formulas if the modal interpretation is to accord with mathematical practice is for their truth-values to be stage-invariant. Set theory is not so parochial that the truth of its theorems varies according to which stage we happen to have ascended to (else we would expect set theoretic investigations to begin by ascertaining which sets have already been formed).

Formally, say that a formula $\psi(\boldsymbol{v})$ (with only the variables $\boldsymbol{v} = v_1, \ldots, v_k$ free) is *stage-invariant* (in symbols: $\textsc{invariant}[\psi]$) if $\Box\forall\boldsymbol{v}(\Box\psi(\boldsymbol{v}) \vee \Box\neg\psi(\boldsymbol{v}))$.[28] The invariance of modalised $\mathcal{L}_\in$-formulas may then be established in the logic of stage theory.

Proposition 4 (Invariance.). $\vdash \textsc{invariant}[\phi^\Diamond]$.

(Henceforth we reserve ϕ and Γ, and their adornments, for non-modal $\mathcal{L}_\in$-formulas and sets of the same, with ψ and Δ for $\mathcal{L}_\in^\Diamond$-formulas and sets of these.)

Proof. We prove $Ev_1, \ldots, Ev_k \vdash \Box\phi^\Diamond \vee \Box\neg\phi^\Diamond$ by induction. For atomic Φ, either $\Diamond\Phi$, so $\Box\Diamond\Phi$ (by E), i.e. $\Box\Phi^\Diamond$, or $\neg\Diamond\Phi$, so $\Box\neg\Phi$, and $\Box\Box\neg\Phi$ (by 4), whence $\Box\neg\Phi^\Diamond$. For $(\neg\phi)^\Diamond$, the result is immediate by the induction hypothesis. For $(\phi_1 \to \phi_2)^\Diamond$, if this formula is true, T and the induction hypothesis yield either (i) $\Box\neg\phi_1^\Diamond$ or (ii) $\Box\phi_2^\Diamond$, in which cases $\Box(\phi_1^\Diamond \to \phi_2^\Diamond)$; or, if the formula is false, (iii) $\Box\phi_1^\Diamond$ and $\Box\neg\phi_2^\Diamond$, so $\Box\neg(\phi_1^\Diamond \to \phi_2^\Diamond)$. For $(\forall v\phi)^\Diamond$, either $\Box\forall v\phi^\Diamond$, so $\Box(\forall v\phi)^\Diamond$ (by 4), or $\neg\Box\forall v\phi^\Diamond$, so $\Diamond\neg\forall v\phi^\Diamond$ and $\Box\Diamond\neg\forall v\phi^\Diamond$ (by E), whence $\Box\neg(\forall v\phi)^\Diamond$. $\Box$

An immediate consequence is that almost all modal distinctions collapse for modalised formulas; indeed the following holds.

[27] The results in the next two subsections draw heavily on Linnebo (2010,n.d.), which proves analogues of propositions 4, 5 and 6 for the (uni-)modal logic he works in.

[28] This strengthens the notion of 'absoluteness' in Parsons (1983) ('stability' in the terminology of Linnebo (n.d.)); in effect, ψ is *stable* if $\Box_\geqslant\forall\boldsymbol{v}(\Box_\geqslant\psi \vee \Box_\geqslant\neg\psi)$. Stability thus ensures that the truth-value of the formula remains invariant at all stages no earlier than the present one when its parameters are formed, but falls short of full invariance.

Proposition 5 (Modal collapse.)**.** The following are provably equivalent in LST on the assumption that $\textsc{invariant}[\psi]$: $\psi, \Box\psi, \Diamond\psi, \Box_{\geqslant}\psi, \Diamond_{\geqslant}\psi, \Box_{>}\psi, \Diamond_{>}\psi, \Box_{\leqslant}\psi, \Diamond_{\leqslant}\psi$. $\Box$

The exceptions $\Box_{<}\psi$ and $\Diamond_{<}\psi$ reflect the fact that at the very first stage, all formulas of the first form are trivially true, and all formulas of the second, trivially false. They are provably equivalent to ψ assuming $\Diamond_{<}\top$.

3.2

A second way we should expect modal theories to be consonant with mathematical practice is by preserving logical relations between formulas. It is plausible to interpret a theory given in $\mathcal{L}_{\in}$ as having modal content only if doing so does not render well-established axioms false, or frequently used patterns of inference invalid. To the extent that set theory is conducted in first-order languages—as it is in the vast majority of cases—we shall see in section 4.6 that all provability and non-provability relations are preserved. Here we show that modalisation preserves provability in classical first-order logic (writing $\Gamma^{\Diamond}$ for $\{\gamma^{\Diamond} \mid \gamma \in \Gamma\}$).

Proposition 6 (Mirroring.)**.** $\Gamma \vdash_{\mathsf{FOL}} \phi$ only if $\Gamma^{\Diamond} \vdash_{\mathsf{LST}} \phi^{\Diamond}$.

Sketch proof. Suppose FOL is axiomatised by adding the axiom Ev to the nonmodal subsystem of LST. An easy induction on FOL-proofs shows that LST proves the modalisation of each FOL axiom and is closed under the modalisation of each FOL rule. For example, when $\Gamma \vdash_{\mathsf{FOL}} \forall v\phi$ is obtained from UG, $\Gamma, Eu \vdash_{\mathsf{FOL}} \phi[u/v]$ has a shorter FOL-proof (satisfying the side conditions on u). By induction hypothesis: $\Gamma^{\Diamond}, \Diamond Eu \vdash_{\mathsf{LST}} \phi^{\Diamond}[u/v]$. Then reason in LST as follows.

$\Gamma^{\Diamond}, Eu \vdash \phi^{\Diamond}[u/v]$	T
$\Gamma^{\Diamond} \vdash \forall v\phi^{\Diamond}$	UG
$\{\Box\gamma^{\Diamond} \mid \gamma \in \Gamma\} \vdash \Box\forall v\phi^{\Diamond}$	propositional LST

The result follows since, for $\gamma \in \Gamma$, $\Gamma^{\Diamond} \vdash \Box\gamma^{\Diamond}$, by props. 4 and 5. $\Box$

4 Bimodal Stage Theory.

Turn now to modal stage theory proper. Let $\mathcal{L}_{\in}^{\Diamond}$ be the bimodal expansion of the language of set theory $\mathcal{L}_{\in}$, the first-order language with $=$ as its only logical predicate and $\in$ as its only non-logical predicate.

In order to formulate a stage theory that upholds Liberalness—the thesis that any sets can form a set—we need first attend to how this thesis should be formulated in $\mathcal{L}^{\Diamond}_{\in}$. In a first-order language, the best we can hope for is a schematic formulation. It should be clear, however, that the inconsistent schema that results simply from modalising the naïve comprehension schema falls wide of the mark.

$$\text{(NCS}^{\Diamond}\text{)} \qquad \Diamond\exists x\Box\forall z(z \in^{\Diamond} x \leftrightarrow \phi^{\Diamond}(z))$$

Quite contrary to the iterative conception, the instance of this schema for $\phi^{\Diamond}(z)$ taken as $z \notin^{\Diamond} z$ says that a set will eventually be formed at some stage that has every satisfier of $z \notin^{\Diamond} z$ *ever* formed at *any* stage as an element. In light of the mirroring result this leads to inconsistency just as inevitably as its nonmodal counterpart does. However this instance is *not* a consequence of Liberalness for the simple reason that no *sets* are ever formed that comprise every future satisfier of $z \notin^{\Diamond} z$. Instead, given *any* non-self-membered sets formed at any stage, *further* non-self-membered sets will be formed at later stages.

The point may be generalised via an $\mathcal{L}^{\Diamond}_{\in}$-regimentation of the notion of indefinite extensibility.[29] Say that a formula $\psi(z)$ is *extensible* if new satisfiers, satisfiers of $\psi(z)$ that were not formed at any earlier stage, will be formed at some later stage; and *indefinitely extensible* if it is always extensible:

$$\textsc{extensible}_z[\psi] =_{\text{df}} \Diamond_{>}\exists z(\psi(z) \wedge \Box_{<}\neg Ez); \quad \text{indext}_z[\psi] =_{\text{df}} \Box\textsc{extensible}_z[\psi].$$

Then, for any indefinitely extensible formula $\psi(z)$, no sets are ever formed that have every future satisfier of $\psi(z)$ among them; for it is always the case that some further satisfier of $\psi(z)$ has yet to be formed. Consequently Liberalness does not imply any instance of NCS$^{\Diamond}$ in which $\psi(z)$ is an indefinitely extensible formula.

In order to formulate Liberalness, then, we need to restrict NCS$^{\Diamond}$ to formulas $\psi(z)$

[29]See, for instance, Michael Dummett (1991). While this regimentation seems to me to capture the kernel of the notion of indefinite extensibility, it differs from Dummett's characterisation in several respects.

Table 5.3: The Axioms of ZF

Extensionality	$\forall x_1 \forall x_2 (\forall z (z \in x_1 \leftrightarrow z \in x_2) \rightarrow x_1 = x_2)$
Foundation	$\forall x \big(\exists z (z \in x) \rightarrow (\exists z \in x)(\forall w \in z)(w \notin x)\big)$
Empty Set	$\exists x \forall z (z \notin x)$
Pairing	$\forall x_1 \forall x_2 \exists y \forall z (z \in y \leftrightarrow z = x_1 \vee z = x_2)$
Union	$\forall x \exists y \forall z (z \in y \leftrightarrow (\exists w \in x)(z \in w))$
Separation	$\forall x \exists y \forall z (z \in y \leftrightarrow z \in x \wedge \phi(z))$
Powerset	$\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$
Infinity	$\exists y ((\exists x \in y) \forall z (z \notin x) \wedge (\forall z \in y)(\exists w \in y)(w = z^+))$
Replacement	$\mathrm{Fn}_{w,z}[\phi] \rightarrow \forall x \exists y \forall z (z \in y \leftrightarrow (\exists w \in x)(\phi(w, z)))$

that 'exactly circumscribe' some sets, in the sense that it is always the case that every satisfier of $\psi(z)$ (and no other set) will be one of them. For every future satisfier of $\psi(z)$ to be one of them, it is necessary, as we just saw, that $\psi(z)$ not be indefinitely extensible. And in fact, provided that $\psi(z)$ is an invariant formula, this condition is also sufficient for $\psi(z)$ to exactly circumscribe some sets. For, at any stage, an invariant formula $\psi(z)$ is inextensible ($\neg\textsc{extensible}_z[\psi]$) just in case every set that ever satisfies $\psi(z)$ is formed then;[30] and this holds just in case $\psi(z)$ exactly circumscribes those sets formed at that stage which satisfy $\psi(z)$ then: for any eventual satisfier of $\psi(z)$ is formed at that stage (and satisfies the invariant formula then) and is therefore one of them; and each of them always satisfies $\psi(z)$. We may therefore formulate Liberalness in the bimodal language $\mathcal{L}_{\in}^{\Diamond}$ by restricting NCS$^{\Diamond}$ to invariant formulas that are not indefinitely extensible.

(Lib) $\quad \textsc{invariant}[\psi] \wedge \neg\mathrm{indext}_z[\psi(z)] \rightarrow \Diamond \exists x \Box \forall z (z \in^{\Diamond} x \leftrightarrow \psi(z)).$

4.1

Let S be the $\mathcal{L}_{\in}$-theory comprising the axioms Extensionality, Foundation, Empty Set, Pairing, Union, Separation, Powerset and Infinity in table 5.3. Zermelo set theory Z adds the axioms of Infinity to S; and Zermelo-Fraenkel set theory, ZF, adds the axiom schema Replacement to Z.[31] The bimodal stage theory—dubbed MST—consists of

[30] The left-to-right direction relies on the well-foundedness of the earlier-later relation.

[31] Notation: $x \subseteq y =_{\mathrm{df}} \forall z (z \in x \rightarrow z \in y)$; $x = y^+ =_{\mathrm{df}} \forall z (z \in y \leftrightarrow z \in x \vee z = x)$; and $\mathrm{Fn}_{w,z}[\phi] =_{\mathrm{df}} \forall z \forall w_1 \forall w_2 (\phi(z, w_1) \wedge \phi(z, w_2) \rightarrow w_1 = w_2)$.

three axioms. The first is simply the modalisation of the Extensionality axiom:

(Ext$^\Diamond$) $$\Box\forall z(z \in^\Diamond x_1 \leftrightarrow z \in^\Diamond x_2) \to x_1 =^\Diamond x_2.$$

The other two axioms provide $\mathcal{L}_\in^\Diamond$-formalisations of the principles of Priority and Plenitude articulated in the iterative conception. These three axioms permit the derivation of the modalised axioms of Zermelo set theory less Infinity, which we label $\mathsf{S}^\Diamond$. With the addition of a further axiom, which elaborates on the iterative conception by telling us something more about its extent, we may also derive the axioms of Infinity and Replacement, thereby recovering all the modalised axioms of Zermelo-Fraenkel set theory, $\mathsf{ZF}^\Diamond$. We consider each axiom in turn.

4.2

The Priority principle has it that, for any set formed at any stage, its elements were all formed at an earlier stage. Since a set is an element of set x just in case it satisfies $z \in^\Diamond x$ (and this formula is invariant), the elements of x are all formed at a stage just in case $z \in^\Diamond x$ is inextensible then. Consequently, the principle of Priority may be formalised in the bimodal language $\mathcal{L}_\in^\Diamond$ as follows:

(Pri) $$\Box\forall x \Diamond_< \neg\textsc{extensible}_z[z \in^\Diamond x].$$

An immediate consequence is that $z \in^\Diamond x$ is inextensible whenever x is formed. The Priority of each element to its set follows from a lemma about extensibility.

Proposition 7. $\mathrm{Pri}, Ex \vdash \neg\textsc{extensible}_z[z \in^\Diamond x]$. □

Lemma 8. $\textsc{invariant}[\psi], \neg\textsc{extensible}_z[\psi] \vdash \psi(z) \to Ez$.

Proof. Suppose $\textsc{invariant}[\psi]$, $\neg\textsc{extensible}_z[\psi(z)]$ and $\psi(z)$. Now $\Diamond Ez$; so, applying WF, we have either (i) $\Diamond_{\leqslant}(Ez \wedge \Box_{>}\neg Ez)$ or (ii) $\Diamond_{>}(Ez \wedge \Box_{<}\neg Ez)$. Since $\Box_{>}\psi(z)$ (by modal collapse, prop. 5) the second case is ruled out as $\neg\textsc{extensible}_z[\psi(z)]$ and so $\Box_{>}\forall z(\psi(z) \to \Diamond_{<}Ez)$. Consequently, (i) holds, whence $\Diamond_{\leqslant}Ez$; therefore Ez (by the forwards necessity of existence). $\Box$

Proposition 9. $\mathrm{Pri}, Ex \vdash z \in^{\Diamond} x \to \Diamond_{<}Ez$. $\Box$

Recall that Regularity is the theorem (schema) of ZF, among whose instances is the axiom Foundation, that states that whenever some set satisfies ϕ, there is some set minimal with respect to $\in$ that satisfies ϕ:

(Regularity) $$\exists x\phi(x) \to \exists x\big(\phi(x) \wedge (\forall z \in x)\neg\phi(z)\big).$$

A pleasing feature of MST is that the modalisation of this schema may be proved on the basis of the Priority axiom. For suppose at some stage $\exists x\phi^{\Diamond}(x)$; then, since the earlier-later relation is well-founded, there is some earliest stage s_0 at which $\exists x\phi^{\Diamond}(x)$. So suppose x is formed at that stage and satisfies $\phi^{\Diamond}$. By the preceding proposition any element z of x must have been formed at some stage before s_0, and since $\forall z\neg\phi^{\Diamond}(z)$ holds at such a stage (by the minimality of s_0), it follows that $\neg\phi^{\Diamond}(z)$; consequently x is an $\in^{\Diamond}$-minimal satisfier of ψ. Formalising this argument in modal stage theory yields the modalisation of Regularity (via a restricted version of the Barcan formula for inextensible formulas).

Proposition 10. $\mathrm{Pri} \vdash \mathrm{Regularity}^{\Diamond}$.

Note first that when an invariant formula ψ_1 is inextensible, its extension does not increase as we move forwards or backwards; and so the forwards, backwards and absolute versions of the Barcan formula may be derived for quantifiers restricted to ψ_1.

Lemma 11. When L is any of $\Box_{<}$, $\Box_{\leqslant}$, $\Box_{>}$, $\Box_{\geqslant}$ or $\Box$:

$$\textsc{invariant}[\psi_1], \neg\textsc{extensible}_v[\psi_1] \vdash \forall v(\psi_1 \to L\psi_2) \to L\forall v(\psi_1 \to \psi_2). \qquad \Box$$

Proof of prop. 10. Assuming Pri, reason in LST as follows. (Here, and throughout, MST axioms stated as assumptions in the result are left tacit in displayed proofs.)

$$
\begin{array}{rll}
\Box_{<}\forall z\neg\phi^{\Diamond}(z), Ex & \vdash z \in^{\Diamond} x \rightarrow \Diamond_{<} Ez & \text{prop. 9} \\
& \vdash z \in^{\Diamond} x \rightarrow \Diamond_{\leqslant}\neg\phi^{\Diamond}(z) & \mathsf{LST} \\
& \vdash z \in^{\Diamond} x \rightarrow \Box\neg\phi^{\Diamond}(z) & \text{modal collapse} \\
& \vdash (\forall z \in^{\Diamond} x)\Box\neg\phi^{\Diamond}(z) & \text{generalisation on } z \\
& \vdash \Box(\forall z \in^{\Diamond} x)\neg\phi^{\Diamond}(z) & \text{lem. 11, prop. 7}
\end{array}
$$

It follows that the consequent of Regularity$^{\Diamond}$ may be proved from Pri on the assumption that $\Diamond(\exists x \phi^{\Diamond}(x) \wedge \Box_{<}\forall z \neg\phi^{\Diamond}(z))$. The result follows by an application of WF. $\Box$

4.3

The third, and final, axiom of bimodal stage theory expresses the Plenitude principle: any sets ever formed will form a set at every later stage. This axiom may be formalised in $\mathcal{L}^{\Diamond}_{\in}$ by a schema along similar lines to Liberalness. When every set that ever satisfies an invariant formula $\psi(z)$ is formed, that is, when $\psi(z)$ is no longer extensible, the set of satisfiers of $\psi(z)$ is formed at every later stage:

(Plen) $$\textsc{invariant}[\psi] \wedge \neg\textsc{extensible}_z[\psi(z)] \rightarrow \Box_{>}\exists x \Box \forall z (z \in^{\Diamond} x \leftrightarrow \psi(z)).$$

Each instance of the Liberalness schema is an immediate consequence of this schema, and the fact that there is no last stage, expressed in D.

Proposition 12. Plen $\vdash$ Lib. $\Box$

Together with the Priority axiom, Liberalness suffices to prove all the remaining axioms of $\mathsf{S}^{\Diamond}$ save for the Powerset axiom.

Proposition 13.

a) Lib $\vdash$ Empty Set$^{\Diamond}$

b) Lib $\vdash$ Pairing$^{\Diamond}$

c) Lib, Pri $\vdash$ Union$^{\Diamond}$

d) Lib, Pri $\vdash$ Separation$^{\Diamond}$

Table 5.4: Defining Conditions of Sets

Axiom	*Set*	*Defining condition*
Empty Set$^\Diamond$	$\varnothing$	$z \neq^\Diamond z$
Pairing$^\Diamond$	$\{x, y\}$	$z =^\Diamond x \vee z =^\Diamond y$
Union$^\Diamond$	$\cup x$	$\Diamond(\exists w \in^\Diamond x)(z \in^\Diamond w)$
Separation$^\Diamond$	$\{z \in^\Diamond x \mid \phi^\Diamond(z)\}$	$z \in^\Diamond x \wedge \phi^\Diamond(z)$

In each case, to apply this schema to show that the set $\{t : \phi^\Diamond(t)\}$, which comprises every satisfier of $\phi^\Diamond$ ever formed, is formed at some stage, it suffices to show that its defining condition $\phi^\Diamond$ is not indefinitely extensible (see table 5.4). This is readily achieved on the basis of some further results about inextensibility.

Lemma 14. Let $\Delta = \{\delta_1, \ldots, \delta_k\}$ be a set of $\mathcal{L}_\in^\Diamond$-formulas without v free. If $\Delta \vdash$ $\textsc{invariant}[\psi]$ and $\Delta \vdash \psi(v) \to Ev$, then $\Delta \vdash \neg\textsc{extensible}_v[\psi(v)]$.

Proof. Suppose $\Delta \vdash \psi(v) \to Ev$ (and w.l.o.g. $\top \in \Delta$).

$$
\begin{array}{rll}
\Diamond_<(\delta_1 \wedge \ldots \wedge \delta_k) \vdash (\Box_< \psi(v) \to \Diamond_< Ev) & & \text{propositional } \mathsf{LST} \\
\textsc{invariant}[\psi], \Diamond_<(\delta_1 \wedge \ldots \wedge \delta_k) \vdash (\psi(v) \to \Diamond_< Ev) & & \text{modal collapse} \\
\textsc{invariant}[\psi], \Diamond_<(\delta_1 \wedge \ldots \wedge \delta_k) \vdash \forall v(\psi(v) \to \Diamond_< Ev) & & \text{UG} \\
\Box_> \textsc{invariant}[\psi], \Box_> \Diamond_<(\delta_1 \wedge \ldots \wedge \delta_k) \vdash \Box_> \forall v(\psi(v) \to \Diamond_< Ev) & & \text{propositional } \mathsf{LST}
\end{array}
$$

The result follows since $\Delta \vdash \Box_> \textsc{invariant}[\psi]$ and $\Delta \vdash \Box_> \Diamond_< \delta$, for $\delta \in \Delta$. □

Proposition 15.

a) $\vdash \neg\textsc{extensible}_z[z \neq^\Diamond z]$

b) $Ex \vdash \neg\textsc{extensible}_z[z =^\Diamond x]$

c) $\neg\textsc{extensible}_t[\psi_1] \vdash \neg\textsc{extensible}_t[\psi_1 \wedge \psi_2]$

d) $\neg\textsc{extensible}_t[\psi_1], \neg\textsc{extensible}_t[\psi_2] \vdash \neg\textsc{extensible}_t[\psi_1 \vee \psi_2]$. □

Proof of prop. 13. Empty Set$^\Diamond$ and Pairing$^\Diamond$ are immediate from prop. 15, Separation$^\Diamond$, from props. 7 and 15. For Union$^\Diamond$, reason as follows, noting that $Ex \vdash w \in^\Diamond x \to Ew$, and $Ew \vdash z \in^\Diamond w \to Ez$.

$$
\begin{array}{rll}
Ex \vdash & w \in^\Diamond x \wedge z \in^\Diamond w \to Ez & \\
\vdash & (\exists w \in^\Diamond x)(z \in^\Diamond w) \to Ez & \\
\vdash & \Diamond(\exists w \in^\Diamond x)(z \in^\Diamond w) \to Ez & \text{lem. 11, invariance}
\end{array}
$$

Applying lem. 14, this yields $Ex \vdash \neg\textsc{extensible}_z[\Diamond(\exists w \in^\Diamond x)(z \in^\Diamond w)]$. □

The need for an axiom stronger than the Liberalness schema, however, becomes apparent when we come to the Powerset axiom. Consider, for instance, ω, the set of finite ordinals. When ω is formed at a stage, the Priority axiom ensures that its elements $0, 1, 2 \ldots$ were formed earlier; and consequently, the elements of any subset y of ω are formed then. The Liberalness schema then ensures that y will eventually be formed at some stage. But it fails to guarantee that there will be some *one* stage when all subsets of ω are formed.[32] By contrast the Plenitude axiom ensures that each subset is formed at the very first stage after the elements of ω are formed, and therefore that all subsets of ω are then formed, and its Powerset formed thereafter. Generalising this argument in bimodal stage theory we obtain the following.

Proposition 16. $\mathsf{MST} \vdash \text{Powerset}^{\Diamond}$.

Proof. Note that $y \subseteq^{\Diamond} x \vdash \neg\text{EXTENSIBLE}_z[z \in^{\Diamond} x] \to \neg\text{EXTENSIBLE}_z[z \in^{\Diamond} y]$ and reason as follows.

$$\begin{array}{rll} y \subseteq^{\Diamond} x, Ex & \vdash \Diamond_{<}\neg\text{EXTENSIBLE}_z[z \in^{\Diamond} x] & \text{Pri} \\ & \vdash \Diamond_{<}\neg\text{EXTENSIBLE}_z[z \in^{\Diamond} y] & \text{propositional } \mathsf{LST} \\ & \vdash \Diamond_{<}\Box_{>}\exists y'\Box\forall z(z \in^{\Diamond} y' \leftrightarrow z \in^{\Diamond} y) & \text{Plen} \\ & \vdash Ey & \text{Ext}^{\Diamond} \end{array}$$

The result follows by DT, lem. 14 and Lib. $\square$

4.4

MST recovers Zermelo set theory less Infinity. But without Replacement, Zermelo set theory, even with Infinity, is well known to be surprisingly weak. Replacement gives us further height, of course, allowing us to attain sets higher up the hierarchy. But this axiom also plays an important role in ensuring that the hierarchy of sets is as wide as it should be at early stages. Given Foundation, recall that the *rank* $\rho(x)$ of a set x may be defined recursively as the least ordinal greater than the rank of each element of x. Equivalently in ZF, $\rho(x)$ is the least ordinal such that $x \subseteq V_{\rho(x)}$.[33] Without Replacement, we are unable to show that the set-theoretic universe is closed under set-theoretic operations that preserve rank. We are unable to show that each set is a

[32] cf. Parsons (1983), pp. 320–1.
[33] See, for instance, Jech (2003), ch. 6

subset of some V_α; in fact, we are unable even to show that each set x is a subset of a transitive set.[34] The universe has not been shown to be closed under rank-preserving operations such as $x \mapsto V_{\rho(x)}$ or $x \mapsto \mathrm{tcl}(x)$, where $\mathrm{tcl}(x)$ is the transitive closure of x, the intersection of all transitive supersets of x.

One way to overcome this deficiency is to make further assumptions about the height of the hierarchy. In the next section we will see how such an assumption permits us to recover Replacement and Infinity. This would also enable us to fill in the gaps in the hierarchy at low rank. In the context of the iterative conception, however, such an approach is unsatisfying. For on this conception, the width of the hierarchy at early stages is not determined by how far the iteration continues: *every* set of rank $\rho(x)$ is formed along side x for the first time at stage $\rho(x)+1$. In particular, $V_{\rho(x)}$ and $\mathrm{tcl}(x)$ are formed whenever x is formed. Consequently, on the this conception, we should expect the set-theoretic universe to be closed under rank-preserving operations without the need for strong assumptions about its extent.

In the non-modal setting, Gabriel Uzquiano has demonstrated how the potential narrowness of Z may be overcome with the addition of a further axiom.[35] As he shows, we may characterise the V_α's without using Replacement, by defining the predicate $R(x,\alpha)$ to hold of set x and ordinal α just in case there is a function f which satisfies the condition:[36]

$$(\star) \qquad Dom(f) = \alpha + 1 \wedge f(\alpha) = x \\ \wedge\ (\forall \beta \leqslant \alpha)\forall y\big(y \in f(\beta) \leftrightarrow (\exists \gamma < \beta)(y \subseteq f(\gamma))\big).$$

Informally $R(x,\alpha)$ tells us that $x = V_\alpha$. Zermelo set theory may then be supplemented by the following axiom, stating that every set is a subset of some cumulative rank, and

[34] See, for instance Mathias (2001).
[35] See Uzquiano (1999).
[36] Von Neumann ordinals may be defined as usual as transitive sets whose elements are well-ordered by $\in$. We reserves $\alpha, \beta, \ldots$ for these ordinals. And define $\alpha < \beta =_{\mathrm{df}} \alpha \in \beta$ and $\alpha \leqslant \beta =_{\mathrm{df}} \alpha < \beta \vee \alpha = \beta$.

permitting us to recover the V_α-hierarchy:

$$\text{(U)} \qquad \forall x \exists y \exists \alpha (R(y, \alpha) \wedge x \subseteq y).$$

On this basis it is readily seen that the set theoretic universe is closed under $x \mapsto V_{\rho(x)}$ or $x \mapsto \operatorname{tcl}(x)$.

Another pleasing feature of bimodal stage theory MST is that it suffices to prove the modalisation of this axiom *without* the need for supplementation. Say that a (future) set M is the *current universe set* if its elements are precisely the sets formed at the current stage of the hierarchy; or equivalently, if every set formed at this stage will be an element of M and the formula $z \in^\Diamond M$ is inextensible. A set M is said to be a *universe set* if it is the current universe set at some stage:[37]

$$\begin{aligned} \textsc{cur-uni}(M) &=_{\text{df}} \forall z (z \in^\Diamond M) \wedge \neg\textsc{extensible}_z[z \in^\Diamond M] \\ \textsc{universe}(M) &=_{\text{df}} \Diamond \textsc{cur-uni}(M). \end{aligned}$$

Notice first that the definition immediately ensures that the current universe set does not exist yet but will exist at every later stage.

Proposition 17. $\mathsf{MST}, \textsc{cur-uni}(M) \vdash \neg EM \wedge \Box_{>} EM$ $\Box$

The following lemma permits us to show that universe sets are transitive and that their elements are the subsets of earlier universe sets. This enables us to show that every universe set is a rank V_α, from which the modalisation of Uzquiano's axioms, $\mathrm{U}^\Diamond$, is an easy corollary.

Lemma 18. Let ψ be a formula without z free. Let $F_\psi(z) =_{\text{df}} \Diamond(Ez \wedge \psi \wedge \Box_{<} \neg\psi)$.

(a) $\mathsf{MST}, \Box\forall z(z \in^\Diamond M \leftrightarrow F_\psi(z)), (\psi \wedge \Box_{<}\neg\psi) \vdash \textsc{cur-uni}(M)$

(b) $\mathsf{MST}, \Box\forall z(z \in^\Diamond M \leftrightarrow F_\psi(z)), \Diamond\psi, \textsc{cur-uni}(M) \vdash (\psi \wedge \Box_{<}\neg\psi)$

(c) $\mathsf{MST}, \Diamond\psi, \vdash \Diamond\exists M(\textsc{universe}(M) \wedge \Box\forall z(z \in^\Diamond M \leftrightarrow F_\psi(z)))$

[37] We reserve the variable M (and its adornments) for universe sets.

Proof. (a) Suppose ψ, then $F_\psi(z)$ implies either $\Diamond_>(Ez \wedge \psi \wedge \Box_< \neg\psi)$, which is ruled out by assumption, or $\Diamond_\leqslant(Ez \wedge \psi \wedge \Box_< \neg\psi_0)$, whence Ez. Then reason as follows.

$$\begin{array}{rll} \psi & \vdash F_\psi(z) \to Ez & \text{just proved} \\ & \vdash \neg\text{EXTENSIBLE}_z[F_\psi(z)] & \text{lem. 14} \end{array}$$

Also, clearly $\psi \wedge \Box_< \neg\psi \vdash \forall z F_\psi(z)$; thus $\psi \wedge \Box_< \neg\psi \vdash \forall z F_\psi(z) \wedge \neg\text{EXTENSIBLE}_z[F_\psi(z)]$. The desired result follows from prop. 3

(b) Note that $\Diamond\psi, \neg(\psi \wedge \Box_< \neg\psi) \vdash \Diamond_<(\psi \wedge \Box_< \neg\psi) \vee \Diamond_>(\psi \wedge \Box_< \neg\psi)$. This is readily proved via WF. So consider each disjunct of the conclusion in turn.

$$\begin{array}{rll} \Box\forall z(z \in^\Diamond M \leftrightarrow F_\psi(z)), \Diamond_<(\psi \wedge \Box_< \neg\psi) & \vdash \Diamond_< \text{CUR-UNI}(M) & \text{by (a)} \\ & \vdash \Diamond_< \Box_> EM & \text{prop. 17} \\ & \vdash EM & \mathsf{LST} \\ & \vdash \neg\text{CUR-UNI}(M) & \text{prop. 17} \end{array}$$

$$\begin{array}{rll} \Box\forall z(z \in^\Diamond M \leftrightarrow F_\psi(z)), \Diamond_>(\psi \wedge \Box_< \neg\psi) & \vdash \Diamond_> \text{CUR-UNI}(M) & \text{by (a)} \\ & \vdash \Diamond_> \neg EM & \text{prop. 17} \\ & \vdash \neg\Box_> EM & \mathsf{LST} \\ & \vdash \neg\text{CUR-UNI}(M) & \text{prop. 17} \end{array}$$

Putting the proofs together yields the result.

(c) follows from (a), Lib and WF ☐

Proposition 19. $\mathsf{MST}, \text{UNIVERSE}(M), x \in^\Diamond M \vdash x \subseteq^\Diamond M$.

Proof.
$$\begin{array}{rll} \forall z(z \in^\Diamond M) \wedge \neg\text{EXTENSIBLE}_z[z \in^\Diamond M], x \in^\Diamond M, z \in^\Diamond x & \vdash Ex & \text{lem. 8} \\ & \vdash Ez & \text{prop. 9} \\ & \vdash z \in^\Diamond M & \end{array}$$

☐

Proposition 20. $\mathsf{MST}, \text{UNIVERSE}(M), x \in^\Diamond M \vdash (\Diamond\exists M' \in^\Diamond M)(\text{UNIVERSE}(M') \wedge x \subseteq^\Diamond M')$.

Proof. Let $\psi =_{\text{df}} \neg\text{EXTENSIBLE}_z[z \in^\Diamond x]$. Assume $\Box\forall z(z \in^\Diamond M' \leftrightarrow F_\psi(z))$ (tacit below).

$$\begin{array}{rll} \text{CUR-UNI}(M), x \in^\Diamond M & \vdash Ex & \text{def. CUR-UNI}(M) \\ & \vdash \Diamond_<(\psi \wedge \Box_< \neg\psi) & \text{Pri, WF} \\ \text{CUR-UNI}(M), z \in^\Diamond x, x \in^\Diamond M & \vdash \Diamond_<(\psi \wedge \Box_< \neg\psi \wedge Ez) & \text{prop. 7} \\ & \vdash \Diamond_<(\text{CUR-UNI}(M') \wedge Ez) & \text{lem. 18(a)} \\ & \vdash \Diamond_<(\text{CUR-UNI}(M') \wedge z \in^\Diamond M') & \text{def. CUR-UNI}(M') \\ & \vdash EM' \wedge \text{UNIVERSE}(M') \wedge z \in^\Diamond M' & \text{prop. 17} \\ & \vdash M' \in^\Diamond M \wedge \text{UNIVERSE}(M') \wedge z \in^\Diamond M' & \text{def. CUR-UNI}(M) \\ \text{UNIVERSE}(M), x \in^\Diamond M & \vdash M' \in^\Diamond M \wedge \text{UNIVERSE}(M') \wedge x \subseteq^\Diamond M' & \mathsf{LST} \end{array}$$

So: $\Diamond\exists M'\Box\forall z(z \in^\Diamond M' \leftrightarrow F_\psi(z)), \text{UNIVERSE}(M), x \in^\Diamond M \vdash (\Diamond\exists M' \in^\Diamond M)(\text{UNIVERSE}(M') \wedge x \subseteq^\Diamond M')$. The first premise may be dropped by lem. 18 since $\Diamond\psi$. ☐

Proposition 21. $\mathsf{MST} \vdash \Box\forall M \Diamond\exists\alpha(M =^{\Diamond} V_{\alpha})$.

Sketch proof. Applying Regularity, the theorem is proved by an $\in$-induction, on the basis of lems. 19 and 20. The base case is straightforward; and when for each $M' \in M$ we have an ordinal $\alpha_{M'}$ and function $f_{M'}$ which satisfy the modalisation of condition ($\star$) (on p. 179) it can be shown that M together with the ordinal $\alpha_M =^{\Diamond} \bigcup_{M' \in M} \alpha_{M'}$ and the function $f_M =^{\Diamond} \left(\bigcup_{M' \in M} f'_M\right) \cup \{\langle \alpha_M, M\rangle\}$ also satisfy this condition. □

Corollary 22. $\mathsf{MST} \vdash \Box\forall x \Diamond\exists y \Diamond\exists\alpha(y =^{\Diamond} V_{\alpha} \wedge x \subseteq^{\Diamond} y)$.

Proof. $\mathsf{MST} \vdash \Diamond\exists M(x \subseteq^{\Diamond} M)$ by the proof of claim (i) in prop. 20. The result then follows from the preceding proposition. □

4.5

While the three axioms of bimodal stage theory thus characterise the shape of the hierarchy, ensuring it receives its intended width, they say little about its extent. MST proves neither the axiom of Infinity nor the axiom of Replacement.[38] Although the former uncontroversially follows from the informal sketch of the iterative conception—according to which ω is formed at every stage after the first limit stage—the latter does not. Nonetheless the modal setting provides a natural means of *elaborating* the iterative conception so as to recover both axioms. We adjoin bimodal stage theory with the following modal reflection schema:[39]

(Refl) $$\Box\Diamond_{>}\forall \boldsymbol{v}(\phi^{\Diamond}(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v}))$$

where ϕ has only the variables $\boldsymbol{v} = v_1, \ldots, v_k$ free.

The thought here is that the hierarchy of sets extends far enough to elude characterisation by any $\mathcal{L}_{\in}$-formula. When the truth of such a formula ϕ at a stage expresses that the hierarchy up to that stage has a certain feature, the truth of its modalisation $\phi^{\Diamond}$ expresses that the entire hierarchy has the corresponding feature. Reflection ensures that we never reach a stage when such features of the hierarchy are not reflected in one of its later stages: it is always the case that there will be a stage when (relative to any

[38] Consider the Kripke model $\langle \omega, \in|_{\omega}, \{V_n\}_{n\in\omega}, \{\in|_{V_n}\}_{n\in\omega}\rangle$.

[39] A similar reflection principle is deployed in Linnebo (n.d.). Thanks to Øystein Linnebo and Sam Roberts for impressing upon me the importance of 'complete' reflection in the present setting.

parameters $\boldsymbol{v}$ formed then) the entire hierarchy has the feature expressed by $\phi^{\Diamond}(\boldsymbol{v})$ just in case the present stage has the feature expressed by $\phi(\boldsymbol{v})$.

In the context of modal stage theory, the modalisation of a more familiar, set-theoretic formulation of reflection may be recovered from Refl:[40]

Proposition 23. $\mathsf{MST}, \text{Refl} \vdash [\forall\alpha(\exists\lambda > \alpha)\exists U(R(U,\lambda) \wedge (\forall \boldsymbol{v} \in U)(\phi(\boldsymbol{v}) \leftrightarrow \phi^{U}(\boldsymbol{v})))]^{\Diamond}$

The result follows from the lemma that, when U is the current stage universe, for parameters formed then, $\phi(\boldsymbol{v})$ is equivalent to the modalisation of it relativisation to U, proved by a straightforward induction.

Lemma 24. $\mathsf{MST}, \textsc{cur-uni}(U), E\boldsymbol{v} \vdash \phi(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})^{U^{\Diamond}}$ $\square$

Proof of prop. 23. Let ψ be the formula $E\alpha \wedge \forall \boldsymbol{v}(\phi^{\Diamond}(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v}))$, and F_{ψ} be as in lem. 18. Assume $\Box\forall x(x \in^{\Diamond} U \leftrightarrow F_{\psi}(x))$ (left tacit below).

$$
\begin{array}{rll}
\textsc{cur-uni}(U), \Diamond\psi \vdash & E\alpha \wedge \forall \boldsymbol{v}\big(\phi^{\Diamond}(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})\big) & \text{lem. 18(b)} \\
\vdash & E\alpha \wedge \forall \boldsymbol{v}\big(\phi^{\Diamond}(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})^{U^{\Diamond}}\big) & \text{lem. 24} \\
\vdash & [\alpha \in U \wedge (\forall \boldsymbol{v} \in U)\big(\phi(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})^{U}\big)]^{\Diamond} & \text{def. } \textsc{cur-uni}(U)\text{, lem. 11} \\
\psi \wedge \Box_{\leqslant}\neg\psi \vdash & \textsc{cur-uni}(U) \wedge [\alpha \in U \wedge (\forall \boldsymbol{v} \in U)\big(\phi(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})^{U}\big)]^{\Diamond} & \text{lem. 18} \\
\Diamond\psi \vdash & \textsc{universe}(U) \wedge [\alpha \in U \wedge (\forall \boldsymbol{v} \in U)\big(\phi(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})^{U}\big)]^{\Diamond} & \text{WF} \\
\vdash & [\exists\beta R(U,\beta) \wedge \alpha \in U \wedge (\forall \boldsymbol{v} \in U)\big(\phi(\boldsymbol{v}) \leftrightarrow \phi(\boldsymbol{v})^{U}\big)]^{\Diamond} & \text{prop. 21}
\end{array}
$$

Consequently, the modalisation of

$$(\dagger) \qquad \forall\alpha(\exists\beta > \alpha)\exists U(R(U,\beta) \wedge (\forall \boldsymbol{v} \in U)(\phi(\boldsymbol{v}) \leftrightarrow \phi^{U}(\boldsymbol{v})))$$

follows from $\Diamond\exists U\Box\forall x(x \in^{\Diamond} U \leftrightarrow F_{\psi}(U))$ and $\Diamond\psi$ (since $\alpha \in V_{\beta}$ iff $\alpha < \beta$). These in turn follow from lem. 18(c) and Refl. Applying a trick of Levy's (1960) we may replace the arbitrary ordinal β in (†) with a limit ordinal λ. The instance of (†) for $(u = \varnothing \wedge \phi(\boldsymbol{v})) \vee (u = \{\varnothing\} \wedge \forall\gamma\exists\delta(\gamma \in \delta))$ permits us to derive the λ-version for $\phi(\boldsymbol{v})$ in S. The result follows by Mirroring. $\square$

The axioms of Infinity and Replacement are recovered by applying mirroring to their familiar, nonmodal derivations from the set-theoretic reflection principle.[41]

Proposition 25.

a) $\mathsf{MST} + \text{Refl} \vdash \text{Infinity}^{\Diamond}$

b) $\mathsf{MST} + \text{Refl} \vdash \text{Replacement}^{\Diamond}$ $\square$

[40] Notation: ϕ^{y} the relativisation of a formula ϕ to a set y is defined as usual, by restricting each quantifier to the elements of y. We reserve λ for limit ordinals (ordinals such that $(\forall\alpha < \lambda)(\exists\beta < \lambda)(\alpha < \beta)$)

[41] The theorem is due to Levy (1960). Our formulation of the non-modal reflection principle follows Kanamori (2009), p. 58

4.6

A set theory does not merely consist in its axioms, but also in their consequences. Since MST proves the modalisation of the axioms of $\mathsf{S}+\mathrm{U}$ and $\mathsf{MST}+\mathrm{Refl}$ proves the modalisations of the axioms of ZF, the mirroring result assures us that these respectively prove (in LST) the modalisations of their classical consequences. But we may also establish a converse to this result: modal stage theory and its extension prove the modalisation of a formula only if it is derivable in $\mathsf{S}+\mathrm{U}$ or ZF respectively.

Theorem 26.

(a) $\mathsf{S}+\mathrm{U} \vdash_{\mathsf{FOL}} \phi$ if and only if $\mathsf{MST} \vdash_{\mathsf{LST}} \phi^{\Diamond}$

(b) $\mathsf{ZF} \vdash_{\mathsf{FOL}} \phi$ if and only if $\mathsf{MST}+\mathrm{Refl} \vdash_{\mathsf{LST}} \phi^{\Diamond}$

The right-to-left direction is established by showing that under the following reverse-translation scheme from $\mathcal{L}_{\in}^{\Diamond}$ to $\mathcal{L}_{\in}$, the reverse-translation $[\psi]^\alpha$ of any theorem of MST or $\mathsf{MST}+\mathrm{Refl}$ is a theorem of the corresponding nonmodal set theory; and that, in nonmodal set theory, an $\mathcal{L}_{\in}$-formula ϕ is provably equivalent to the reverse-translation of its modalisation, $[\phi^{\Diamond}]^\alpha$.

$$\begin{aligned}
[x=y]^\alpha &=_{\mathrm{df}} x=y \wedge x \in V_\alpha \wedge y \in V_\alpha \\
[x\in y]^\alpha &=_{\mathrm{df}} x\in y \wedge x \in V_\alpha \wedge y \in V_\alpha \\
[\neg\psi]^\alpha &=_{\mathrm{df}} \neg[\psi]^\alpha \\
[\psi_1 \to \psi_2]^\alpha &=_{\mathrm{df}} [\psi_1]^\alpha \to [\psi_2]^\alpha \\
[\forall x \psi]^\alpha &=_{\mathrm{df}} (\forall x \in V_\alpha)[\psi]^\alpha \\
[\Box_{>}\psi]^\alpha &=_{\mathrm{df}} (\forall \beta > \alpha)[\psi]^\beta \\
[\Box_{<}\psi]^\alpha &=_{\mathrm{df}} (\forall \beta < \alpha)[\psi]^\beta
\end{aligned}$$

Lemma 27.

(a) $\mathsf{MST} \vdash_{\mathsf{LST}} \psi$ only if $\mathsf{S}+\mathrm{U} \vdash_{\mathsf{FOL}} [\psi]^\alpha$

(b) $\mathsf{MST}+\mathrm{Refl} \vdash_{\mathsf{LST}} \psi$ only if $\mathsf{ZF} \vdash_{\mathsf{FOL}} [\psi]^\alpha$

(c) $\mathsf{S}+\mathrm{U} \vdash \phi \leftrightarrow [\phi^{\Diamond}]^\alpha$ □

Proof of thm. The left-to-right direction is immediate by props. 10, 13, 16, 22 and 25, and mirroring. For the right-to-left direction of (a): if $\mathsf{MST} \vdash \phi^{\Diamond}$, then $\mathsf{S}+\mathrm{U} \vdash [\phi^{\Diamond}]^{\alpha}$ (by lem. 27(a)) and therefore $\mathsf{S}+\mathrm{U} \vdash \phi$ (by lem 27(c)). Similarly for (b). $\square$

5 Philosophical Assessment.

We have taken the tense in the informal sketch of the iterative conception more seriously than has typically been the case (although not literally). The result is a stage theory MST formulated in the bimodal language $\mathcal{L}_{\in}^{\Diamond}$, whose three axioms provide natural formalisations of the principles of Extensionality, Priority and Plenitude found in the informal sketch of the iterative conception. Against the background of the tense-*like* logic LST set out in section 2, this stage theory recovers a good deal of modalised set theory: MST suffices to prove the modalisation of all theorems of $\mathsf{S}+\mathrm{U}$, Zermelo set theory less Infinity but extended with the resources to prove that every set is a subset of some rank V_α. And with the addition of the modal reflection principle Refl, MST proves the modalisation of all theorems of ZF. Such a modal interpretation of set theory was argued in section 3 to fit well with mathematical practice.

As argued in section 1, the key advantage of modal stage theory over the more familiar nonmodal stage theories is that it permits us to uphold Liberalness. We can allow that any sets *can* form a set, forestalling difficulty questions about what it takes for some sets to form another. But what about its *modal* rivals? The advantage just claimed *is* shared by the unimodal axiomatisations given by Parsons (1983) and Linnebo (n.d.). Working respectively in second-order and plural settings, and taking a single weakly-forwards looking operator as primitive, each is able to interpret ZF on the basis of a unimodal formalisation of Liberalness. All else being equal we should prefer the less ideologically profligate views.[42]

But all else is not equal. In the first place, there is a question over whether Parsons' and Linnebo's theories *do* employ a leaner ideology. For unlike the present axiomatisation, both these theories employ quantificational resources other than singular

[42]Parsons also presents a first-order theory. This theory however uses extra modal resources in the form of a scoping device (see Parsons (1983, appendix 1, §2)) and suffers from similar deficiencies to the second-order theory when it comes to Powerset and Foundation.

first-order quantifiers. It is not a straight choice between one modal operator or two but between a modal operator and higher-order or plural quantification or two modal operators. And, in fact, I think some philosophers will find the latter combination more congenial for a foundational account of the iterative conception of set. For once we have accepted the intelligibility and acceptability of one modal operator—as we must if *either* the unimodal or the bimodal approach is to get off the ground—there can be no serious doubts about adding a second. Considerations of parity suggest there will be a symmetric case for the good standing of the second operator. But for proponents of a Quinean view towards higher-order and plural quantification, who take this to be nothing other than covert quantification over sets, there is a serious concern about the higher-order or plural formulations of Liberalness. Once the set-theoretic import of the quantifiers has been made explicit, Liberalness emerges as the trivial truth that, for any *set*, there can be a set with the same elements. And such a truism can do no serious explanatory work. The plural or higher-order unimodal account presupposes an anti-Quinean stance towards both primitive modality and plural or higher-order quantification whereas MST only requires the former.

Second, even with plural or higher-order resources, a unimodal language leads to a substantial drop in expressive power. Without strictly forwards-looking modal operators we cannot formulate Plen. Parsons and Linnebo each instead deploy an analogue of the weaker Liberalness thesis as an axiom. But as we saw above, the stronger principle plays an important role in the derivation of Powerset. Without the Plenitude axiom, we are unable to prove that the condition *subset of* x is inextensible when x is formed, as we did in the proof of proposition 16. Instead both Parsons and Linnebo resort to positing such inextensibility as an additional axiom to recover Powerset. Moreover, no formula containing only strictly or weakly forwards-looking operators is provably equivalent to the Priority axiom Pri.[43] Nor is any set of such formulas able to characterise the well-

[43] Proof: Given a model $\mathcal{M} = \langle S, <, \{M_s\}_{s\in S}, \{|\in|_s\}_{s\in S}\rangle$, and $r \in S$, define $S_r =_{\mathrm{df}} \{s \in S \mid s \geqslant r\}$ and $\mathcal{M}_r = \langle S_r, <|_{S_r}, \{M_s\}_{s\in S_r}, \{|\in|_s\}_{s\in S_r}\rangle$. Let ψ be a formula containing only forwards-looking modal operators. Then a trivial induction shows that for $s \in S_r$ and σ over $\mathcal{M}_r$:

$$|\psi|_{\mathcal{M}_r,s}{}^{\sigma} = |\psi|_{\mathcal{M},s}{}^{\sigma}$$

Clearly however Pri lacks this property (unless r is the first stage, deleting the hierarchy of founded models as WF does in proposition 1.[44] Instead Parsons and Linnebo each add axioms stating that being an element of a formed set is inextensible. This ensures that elements are formed no later than their sets but does not ensure that elements are formed *prior* to their sets. And without Priority or the well-foundedness of the earlier-later relation, we are unable to recover Foundation as in proposition 10. Instead Parsons and Linnebo each resort simply to adding Foundation as a further axiom.

This provides a clear reason to prefer the bimodal axiomatisation of the iterative conception over its unimodal rivals: the bimodal account has greater explanatory power. Simply to stipulate the inextensibility of the subset relation or the well-foundedness of the element-set relation is to give up on explaining why these properties flow from the core principles behind the iterative conception. By adopting a bimodal language we have the expressive resources to fully articulate these principles, and thereby to prove what we would otherwise have to take as basic.

sets prior to r will falsify Pri at r). Consequently Pri and ψ are not provably equivalent, by soundness.

[44] Applying the result in n. 43 it is easily seen that any formula ψ containing only forwards looking modal operators is valid on the well-founded frame $\langle \mathbb{N}, <, \{M_k\}_{k \in \mathbb{N}} \rangle$ only if it is also valid on the ill-founded frame $\langle \mathbb{Z}, <, \{M_k\}_{k \in \mathbb{Z}} \rangle$ (where each $M_k = \{0\}$).

Part III

Metasemantics

6 How Universes Expand

The characteristic claim of the expansionist relativist is that the universe is always open to expansion. Somehow or other, we can expand the range of even unrestricted quantifiers, adding new objects—in particular, new sets—into the universe. In part I, we saw that this permits the expansionist to avoid many of the most damaging objections against restrictionist relativism, according to which the domain of our quantifiers is always restricted to be a proper subdomain of the entire universe. In part II, we generalised the familiar case against absolutism in favour of relativism. The absolutist is forced to give up highly plausible plenitude principles about sets. Meanwhile the relativist can sustain these and is able, moreover, to embed them within an appealing motivation for standard set theory, via a modal axiomatisation of the iterative conception of set. All else being equal, therefore, we have reasons both to prefer relativism over absolutism and expansionist relativism over its restrictionist counterpart.

But it remains to be seen that all else is equal. In particular, in the case of interpretational expansionism, it remains for the expansionist to give a positive account of her key contention that the lexicon of a language can always be reinterpreted in such a way that the universe of discourse expands. This is all the more pressing, since it is not obvious how universe expansion functions. In the case of the most prominent restrictionist view, contextual restrictionism, we do have well-developed semantic accounts of quantifier domain restriction, even if their details remain controversial.

Correspondingly, the interpretational expansionist needs to advance a well-motivated *meta*semantic account of how the lexicon can be reinterpreted so that new objects come to be added to the universe. Expansionism will not emerge as a viable alternative to absolutism and restrictionism until we have at least the beginnings of an account of how this could happen.

In this final chapter, therefore, I shall sketch the beginnings of an account of universe expansion on behalf of the interpretational expansionist. Against the background of some substantial but independently motivated assumptions from metasemantics, I outline how on the basis of speaking an initial language, a community of speakers might come to so use their language as to expand the initial universe of the language with a further rank of sets. This account will be presented at a high level of idealisation. An account of how universes *in fact* expand—assuming that they do—would require an empirical investigation of natural language that philosophers are ill-placed to conduct from their armchairs. My principal focus will instead be on whether an appropriate pattern of use *could* bring about universe expansion. Accordingly, I shall focus on the plight of imaginary, idealised linguistic communities, whose lexicon, syntax and patterns of use are simple enough to make the account philosophically tractable. But I shall make some attempt to isolate patterns of use that are sufficiently straightforward and have enough consonance with our pre-theoretic inclinations as to not obviously fail to be idealisations of the patterns we in fact follow.

My primary aim will be to give an account of universe expansion from the relativist's point of view. To the extent, however, that the assumptions behind the account are independently attractive this account of universe expansion may also provide fresh considerations in favour of adopting this standpoint. A secondary aim will be to assess the effectiveness of such an argument. In the final section, I shall also suggest that if the account of universe expansion can be construed as an argument for relativism, it pushes the relativist towards a more extreme version of her view. To the extent that the account of universe expansion offers an argument in favour of relativism over absolutism, it also supports relativism about inter-perspective modal operators as much as relativism about quantifiers.

1 Use

1.1 Patterns of use

The account of universe expansion is based on a charitable assumption about how facts about use determine facts about meaning. But before we come to this let us start with use. The core patterns of use that will interest us concern the mental or linguistic behaviour of a given linguistic community, which we may gloss to a first approximation as members of the community finding it incoherent to accept certain propositions and to reject certain others. In particular, we shall be interested in inter-perspective patterns, in which the use of one language relates to the use of another. As in chapter 4, we shall record such patterns of use as labelled sequents.

For example, that English speakers find it incoherent both to accept and reject what is expressed in English by 'Snow is white' is recorded by the *intra*-perspective sequent:

$$\textit{English}\text{: `Snow is white'} \Rightarrow_{\textit{English}} \textit{English}\text{: `Snow is white'}$$

That bilingual English-German speakers find the same sort of incoherence in accepting this proposition whilst rejecting what is expressed in German by 'Schnee ist Weiß' is expressed by the *inter*-perspective sequent

$$\textit{English}\text{: `Snow is white'} \Rightarrow_{\textit{Bilingual}} \textit{German}\text{: `Schnee ist weiß'}$$

Such patterns may also extend to more than two propositions. We write

$$\begin{array}{l}\textit{English}\text{: `That snow is white or grey'} \\ \qquad\qquad \Rightarrow_{\textit{Bilingual}} \textit{English}\text{: `That snow is grey', } \textit{German}\text{: `Das Schnee ist weiß'}\end{array}$$

to indicate that bilingual English-German speakers find it incoherent to accept that the snow is white or grey and to reject both what is expressed in English by 'That snow is grey' and what is expressed in German by 'Das Schnee ist weiß'.

More generally, when Γ and Θ are sets of labelled sentences, $\Gamma \Rightarrow_C \Theta$ may be approximately glossed as saying that, for any assignment σ, members of speech community C find it incoherent both to accept what is expressed by each antecedent sentence, labelled in Γ, under its labelled interpretation and σ and to reject what is expressed by each consequent sentence, labelled in Θ, under its labelled interpretation and σ. As before we call the propositions expressed by antecedent and consequent sentences relative to their labelled perspective antecedent and consequent propositions.[1] Sequent rules of the forms

$$\frac{\Gamma_1 \Rightarrow_C \Theta_1}{\Gamma \Rightarrow_C \Theta} \qquad \frac{\Gamma_1 \Rightarrow_C \Theta_1 \qquad \Gamma_2 \Rightarrow_C \Theta_2}{\Gamma \Rightarrow_C \Theta}$$

simply record material implications, respectively that $\Gamma_1 \Rightarrow_C \Theta_1$ only if $\Gamma \Rightarrow_C \Theta$ and that both $\Gamma_1 \Rightarrow_C \Theta_1$ and $\Gamma_2 \Rightarrow_C \Theta_2$ only if $\Gamma \Rightarrow_C \Theta$.

The labelled sequents—informally glossed 'members of community C find it incoherent to accept... and to reject... '—make a claim about how members of community C *use* language. There is considerable room for disagreement, however, over how these sequent are best understood. While there is widespread agreement that the facts about language use, in particular, facts about inference, have an important role to play in determining facts about meaning, there is rather less agreement about what is to count as 'use' in the relevant sense. For instance, there is room for disagreement about whether use facts include facts about intentional attitudes such as knowledge and belief and about whether use facts are descriptive, describing the speakers linguistic and mental behaviour or normative, concerning their evaluative attitudes towards the same.[2] As far as possible, I would like to remain neutral on such questions. Granted the charitable assumption below, the different readings of labelled sequents amount to different metasemantic theories, offering competing answers to the metasemantic question of how, collectively or

[1]This is to implement an inter-perspective version of Restall's (2005) proposal that sequents be read as incoherence relations.

[2]Davidson (1973), for example, takes the use facts to consist of which sentences speakers 'hold true'; Lewis (1974) has the more ambitious aim of ultimately reducing the use facts to facts about speakers as a physical system. But while both of these seem to be descriptive accounts, Brandom (1998) emphasises normative features of use.

individually, our occurrent or dispositional mental and linguistic behaviour, or normative attitudes towards the same, determine semantic content. But, by and large, the account of universe expansion does not turn on such details. Consequently the relativist can afford to remain neutral on such difficult issues.

The argument does however impose some general constraints on how such labelled sequents may be understood under such an incoherence reading. In particular, we suppose that labelled sequents so interpreted conform to the usual structural rules for Gentzen sequents. And this is incompatible with some use-related readings of the sequents.

First of all, members of community C find (or appear to find) it incoherent both to accept and reject what is expressed by the same sentence.

$$\text{(REP.)} \qquad \varsigma\colon \phi \Rightarrow_C \varsigma\colon \phi$$

Of course, we do not in general want to impute universal freedom from contradiction across the community C. Linguistic communities are often unruly: error is commonplace, contradiction is not unusual, and even appropriate normative attitudes towards contradiction are not universal—including among experts.[3] For such communities, REP cannot be understood as saying that that *every* member of the community, or every expert, always finds it incoherent both to accept and reject what is expressed by a single sentence. Such unruliness however is compatible with a wide variety of weaker readings of the sequent, according to which many or most speakers—or perhaps, many or most members of a privileged segment of the community—often find such an attitude to be incoherent.

Second, we assume that incoherence is never found to be relieved by taking on further commitments. If members of community C find it incoherent to accept some propositions and to reject some other propositions, members of community C also find it incoherent both to do this *and* to accept and/or reject further propositions. When

[3]Consider, for instance, Priest (2006).

$\Gamma_0 \subseteq \Gamma$ and $\Theta_0 \subseteq \Theta$:[4]

(DIL) $$\frac{\Gamma_0 \Rightarrow_C \Theta_0}{\Gamma \Rightarrow_C \Theta}$$

Third, and last, there is always some way to extend an acceptance/rejection distribution that is not found to be incoherent to encompass the acceptance or rejection of a new proposition, without the new distribution being found to be incoherent. If members of C find it incoherent to adopt a certain acceptance/rejection distribution, regardless of whether they also accept a new proposition, or reject it, then members of C find the initial distribution to be incoherent. This is recorded in the following sequent rule.

(CUT) $$\frac{\varsigma\colon\phi, \Gamma \Rightarrow_C \Theta \qquad \Gamma \Rightarrow_C \Theta, \varsigma\colon\phi}{\Gamma \Rightarrow_C \Theta}$$

In conjunction with other patterns of use, the last two assumptions will permit us to derive infinitely many sequents. Given our reading of sequent rules simply as statements of material implication, this means that infinitely many patterns of use are imputed to speakers. This severely diminishes the plausibility of certain use readings of the sequent but still leaves many intact. We cannot plausibly interpret each distinct sequent as imputing a distinct occurrent mental state of finding incoherent. For finite communities of finite beings do not have infinitely many such states. But there is no similar problem of principle in taking 'finding' an acceptance/rejection distribution to be incoherent in dispositional terms. Moreover, no problem arises on a descriptive account according to which the sequent indicates simply that speakers do not both accept the antecedent propositions and reject the consequent ones.

Notice also that conformity to CUT provides a natural link between the pattern of use described by $\Gamma \Rightarrow_C \varsigma\colon\phi$ and members of the community inferring what is expressed by the conclusion ϕ relative to its labelled perspective from what is expressed by the premisses labelled in Γ. Whenever $\Gamma \Rightarrow_C \varsigma\colon\phi$ and members of C do not find it incoherent to accept each of the antecedent propositions expressed by the premisses in Γ (that is, $\Gamma \not\Rightarrow_C$) then adherence to CUT ensures that they do not find it incoherent to accept what

[4]To ease readability, weakening steps in sequent proofs are usually left tacit below.

Table 6.1: The lexicon of $\mathcal{L}^{pl}(\in, \{\cdot\})$

Category e and ee	singular and plural variables: $x, y, \ldots$ and $xx, yy, \ldots$
Category $t/(e, e)$	singular identity predicate: $=$
Category $t/(ee, ee)$	the plural identity predicate: $\equiv$
Category $t/(e, ee)$	the member-'plurality' predicate: $\prec$
Category $t/(e, e)$	the element set predicate: $\in$
Category e/ee	the set-term operator: $\{\cdot\}$
Category ee/t	plural term-forming operators: $\lambda_p x$, $\lambda_p y. \ldots$
Category t/t and $t/(t, t)$	connectives: $\neg$ and $\wedge$
Category t/t	singular and plural quantifiers: $\forall x, \ldots$ and $\forall xx, \ldots$

is expressed by the conclusion at the relevant perspective in addition to the antecedent propositions ($\Gamma, \varsigma\colon \phi \not\Rightarrow_C$). $\Gamma \Rightarrow_C \varsigma\colon \phi$ ensures that the premisses license inference to the conclusion without risk of loss of coherence.[5]

Consequently, the usual structural rules (adapted for labelled sequents) are compatible with a wide range of interpretations of the sequent, $\Gamma \Rightarrow_C \Theta$, as describing what may be roughly glossed as members of C finding it incoherent to accept the antecedent propositions and reject the consequent ones. As a result, derivations in a Gentzen-style sequent calculus provide a useful tool for examining patterns of use recorded in this way. And we shall make free use of such derivations in what follows.

1.2 The Deliberate Extenders

Let us now turn to the idealised linguistic community, whose imaginary attempt to expand their universe will be our principal focus. The extent of the idealisation is immediately apparent in their lexicon. We assume that members of the community speak an interpreted plural first-order language $\mathcal{L}^{pl}(\in, \{\cdot\})$, modifying the lexicon of the language of set theory with urelements. The lexicon is displayed in table 6.1. The set predicate β employed in earlier chapters is replaced by a set-term forming operator $\{\cdot\}$, which given a plural term $\tau\tau$, yields a singular term $\{\tau\tau\}$, which may be read 'the set of $\tau\tau$'. This allows us to recover the old predicate by definition: $\beta x =_{\mathrm{df}} \exists yy(x = \{yy\})$.

The use of logical expressions by members of this community embodies a com-

[5]The points made in the last four paragraphs draw on Restall (2005), who argues that the structural rules can be seen as giving a plausible minimal theory of an incoherence reading of the sequent; and that CUT can be seen as underlying inference.

mitment to the free plural first-order logic FPFO outlined in chapter 4. Let p be the perspective their language corresponds to (in contexts where no restrictions are imposed on their quantifiers). Then members of the community conform to patterns of use corresponding to the p-instances of the sequent rules recorded in table 6.3 (on page 197 below) together with the patterns of use that may be derived from these using the structural rules. For example the rules for $\neg$, allow us to derive.[6]

$$p\colon \phi, p\colon \neg\phi \Rightarrow$$

This records the fact that members of the imaginary community find it incoherent to both accept what is expressed in their language by a sentence ϕ and what is expressed by its negation $\neg\phi$.

We make some limited assumptions about the interpretation of the Extenders' lexicon. First we suppose that the expressions receive their intended interpretation over the initial universe M_p. In particular, the familiar satisfaction clauses for the logical expressions displayed in table 6.2 hold for all assignments σ to singular and plural variables. Notice that these semantics ensure that the initial universe M_p comprises those objects a such that Ev is true relative to p when v is assigned to a (where Ev abbreviates $v = v$, as usual). Second we assume that M_p is 'transitive' in that whenever it contains a set it also contains the elements of that set. Neither of the assumptions should trouble the absolutist. After all he contends that we can come to the intended interpretation of these expressions over the absolute universe. And any universe that is a plausible candidate for absolute will be transitive (otherwise it lacks elements of its members).

For some time this interpretation of the lexicon remains unchanged. After a while, however, a philosophical divide starts to open within the imaginary community. Some of the speakers become staunch relativists about quantifiers; other become devout absolutists. The absolutists become accustomed to patiently explaining to their relativist contemporaries that when they use quantifiers without any contextual restrictions, they succeed in quantifying over everything, really absolutely everything with no restrictions

[6] As in chapter 4, this is short for $\{p\colon \phi, p\colon \neg\phi\} \Rightarrow \varnothing$. We apply similar abbreviations below.

Table 6.2: Intended truth conditions for logical expression relative to p

Let $\text{Den}_\sigma(vv)$ abbreviate 'the (zero or more) objects σ maps vv to'; $\text{Den}_\sigma(\langle\lambda_p v.\phi\rangle)$ abbreviate 'the zero or more members a of M_p such that $\phi[u/v]$ is true relative to p under $\sigma[a/u]$'.

$u = v$ is true relative to p under σ iff $\sigma(u)$ is identical to $\sigma(v)$, which is a member of M_p

$\tau\tau_1 \equiv \tau\tau_2$ is true relative to p under σ iff $\text{Den}_\sigma(\tau\tau_1)$ are identical to $\text{Den}_\sigma(\tau\tau_2)$, which are members of M_p

$u \prec \langle\lambda_p v.\phi\rangle$ is true relative to p under σ iff $\sigma(u)$ is one of $\text{Den}_\sigma(\langle\lambda_p v.\phi\rangle)$.

$\neg\phi$ is true relative to p under σ iff ϕ is not true relative to p under σ.

$\phi_1 \wedge \phi_2$ is true relative to p under σ iff ϕ_1 is true and ϕ_2 is true relative to p under σ.

$\forall v\phi$ is true relative to p under σ iff, for every a in M_p, ϕ is true relative to p under $\sigma[v/a]$.

$\forall vv\phi$ is true relative to p under σ iff, for all zero or more aa in M_p, ϕ is true relative to p under $\sigma[vv/aa]$.

Table 6.3: Patterns of use for logical expressions.

$$\frac{\Gamma \Rightarrow \Theta, \varsigma\colon \phi}{\varsigma\colon \neg\phi, \Gamma \Rightarrow \Theta}\ \text{L}\neg \qquad \frac{\varsigma\colon \phi, \Gamma \Rightarrow \Theta}{\Gamma \Rightarrow \Theta, \varsigma\colon \neg\phi}\ \text{R}\neg$$

$$\frac{\varsigma\colon \phi_1, \varsigma\colon \phi_2, \Gamma \Rightarrow \Theta}{\varsigma\colon \phi_1 \wedge \phi_2, \Gamma \Rightarrow \Theta}\ \text{L}\wedge \qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon \phi_1 \qquad \Gamma \Rightarrow \Theta, \varsigma\colon \phi_2}{\Gamma \Rightarrow \Theta, \varsigma\colon \phi_1 \wedge \phi_2}\ \text{R}\wedge$$

$$\frac{\Gamma \Rightarrow \Theta, \varrho\colon \phi}{\varsigma\colon \tau_1 = \tau_2, \Gamma \Rightarrow \Theta, \varrho\colon \phi(\tau_2/\tau_1)}\ \text{L} = \qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon \Phi(\tau)}{\Gamma \Rightarrow \Theta, \varsigma\colon \tau = \tau}\ \text{R} =$$

$$\frac{\Gamma \Rightarrow \Theta, \varrho\colon \phi}{\varsigma\colon \tau\tau_1 \equiv \tau\tau_2, \Gamma \Rightarrow \Theta, \varrho\colon \phi(\tau\tau_2/\tau\tau_1)}\ \text{L} \equiv \qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon \Phi(\tau\tau)}{\Gamma \Rightarrow \Theta, \varsigma\colon \tau\tau \equiv \tau\tau}\ \text{R} \equiv$$

$$\frac{\varsigma\colon E\tau, \varsigma\colon \phi[\tau/v], \Gamma \Rightarrow \Theta}{\varsigma\colon \tau \prec \langle\lambda_\varsigma\colon v.\phi\rangle, \Gamma \Rightarrow \Theta}\ \text{L}\lambda \qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon \phi[\tau/v] \qquad \Gamma \Rightarrow \Theta, \varsigma\colon E\tau}{\Gamma \Rightarrow \Theta, \varsigma\colon \tau \prec \langle\lambda_\varsigma\colon v.\phi\rangle}\ \text{R}\lambda$$

$$\frac{\varsigma\colon \phi[\tau/v], \Gamma \Rightarrow \Theta \qquad \Gamma \Rightarrow \Theta, E\tau}{\varsigma\colon \forall v\phi, \Gamma \Rightarrow \Theta}\ \text{L}\forall v \qquad \frac{\varsigma\colon Eu, \Gamma \Rightarrow \Theta, \varsigma\colon \phi[u/v]}{\Gamma \Rightarrow \Theta, \varsigma\colon \forall v\phi}\ \text{R}\forall v$$

$$\frac{\varsigma\colon \phi[\tau\tau/vv], \Gamma \Rightarrow \Theta \qquad \Gamma \Rightarrow \Theta, E\tau\tau}{\varsigma\colon \forall vv\phi, \Gamma \Rightarrow \Theta}\ \text{L}\forall vv \qquad \frac{\varsigma\colon Euu, \Gamma \Rightarrow \Theta, \varsigma\colon \phi[uu/vv]}{\Gamma \Rightarrow \Theta, \varsigma\colon \forall vv\phi}\ \text{R}\forall vv$$

$$\frac{}{\Rightarrow \varsigma\colon E\langle\lambda_\varsigma v.\phi\rangle}\ \text{PC}$$

Side constraints: in R$\forall v$ (R$\forall v$) the variable u (uu), does not occur free in the end sequent; in R $=$, the term τ does occur free in $\Phi(\tau)$.

whatsoever. The relativists are well apprised of this use of quantifiers, and do not disagree with the absolutist's contentions, so described, but do think that these miss the point. They think that it remains to be seen whether they might not come to speak a more liberal language, one in which they might describe the expressive limitations of their present language. Tiring of the apparent impasse, they resolve to isolate themselves from the rest of their community, and attempt to do just this. Many years later we return to see how this breakaway linguistic community, now calling themselves the Deliberate Extenders, have gotten on.

The relativist exiles have not forgotten how to speak their old language but seem to have become bilingual. They have attempted to expand their universe by cultivating new patterns of use alongside the old ones. In doing so they have left their initial lexicon unchanged. Instead they rely on contextual clues and conventional mechanisms in order to indicate whether they are speaking in the old language or the new one.

Extenders retain their previous commitment to plural logic in the old language. But they also have a similar commitment in the new language. When speaking the new language their use continues to conform to the same sequent rules for the logical expressions in the lexicon as before, where p is replaced by p', the perspective term for the new language (in unrestricted contexts). Moreover they continue to use their old λ_p-terms in the new language in much the way that they used them in the old language, corresponding to the informal gloss 'the members of the *initial* universe that satisfy ϕ relative to p'. Extenders' use of terms of the form $\langle \lambda_p v.\phi \rangle$ conforms to these patterns.[7]

$$\frac{p\colon E\tau, p\colon \phi[\tau/v], \Gamma \Rightarrow \Theta}{p'\colon \tau \prec \langle \lambda_p v.\phi \rangle, \Gamma \Rightarrow \Theta}\ \text{L}'\lambda_p \qquad \frac{\Gamma \Rightarrow \Theta, p\colon \phi[\tau/v] \qquad \Gamma \Rightarrow \Theta, p\colon E\tau}{\Gamma \Rightarrow \Theta, p'\colon \tau \prec \langle \lambda_p v.\phi \rangle}\ \text{R}'\lambda_p$$

Much like our own commitment to logic, the Extenders commitment to plural logic is open-ended in the sense of McGee (2000) and Williamson (2006). It is not limited to conforming to the patterns of use recorded by instances of these schemas whose substituends are drawn from the current lexicon. Extenders remain committed to these

[7]Since the lexicon remains the same, the new language lacks $\lambda_{p'}$ terms (with a prime on the perspective term) and accordingly the p' instance of the rules for these expressions is redundant.

patterns when they come to learn new expressions. In particular, their use continues to conform to these patterns when side-formulas are labelled by perspective-terms other than p or p'. And their use with the singular and plural identity predicates conforms not merely to the intra-perspective version of substitutivity but also its inter-perspective analogues.

$$\frac{\Gamma \Rightarrow \Theta, \varrho\colon \phi}{\varsigma\colon \tau_1 = \tau_2, \Gamma \Rightarrow \Theta, \varrho\colon \phi(\tau_2/\tau_1)}\ \mathrm{L}= \qquad \frac{\Gamma \Rightarrow \Theta, \varsigma\colon \phi}{\varrho\colon \tau\tau_1 \equiv \tau\tau_2, \Gamma \Rightarrow \Theta, \varrho\colon \phi(\tau\tau_2/\tau\tau_1)}\ \mathrm{L}\equiv$$

In order to make the argument for universe expansion, it suffices to focus on the patterns of use Extenders adopt for certain canonical forms of atomic sentence containing mathematical expressions. Members of the relativist community's use of set-term identity statements embodies an inter-language version of Frege's Basic Law V.[8] Extenders have come to find it incoherent to accept what is expressed in the initial language by the plural identity statement $\tau\tau_1 \equiv \tau\tau_2$ yet reject what is expressed in the new language by the corresponding singular identity statement $\{\tau\tau_1\} \equiv \{\tau\tau_2\}$, and vice versa.

$$p'\colon \{\tau\tau_1\} = \{\tau\tau_2\} \Rightarrow p\colon \tau\tau_1 \equiv \tau\tau_2 \qquad p\colon \tau\tau_1 \equiv \tau\tau_2 \Rightarrow p'\colon \{\tau\tau_1\} = \{\tau\tau_2\}$$

Likewise, acceptance and rejection of propositions expressed in the new language by canonical element-set statements of the form $\tau \in \{\tau\tau\}$ are co-ordinate in the natural way with acceptance and rejection of what is expressed by the corresponding member-'plurality' statements in the initial language. Extenders have come to find it incoherent to accept what is expressed in the old language by $\tau \prec \tau\tau$ and yet to reject what is expressed in the new language by $\tau \in \{\tau\tau\}$, and vice versa.

$$p'\colon \tau \in \{\tau\tau\} \Rightarrow p\colon \tau \prec \tau\tau \qquad p\colon \tau \prec \tau\tau \Rightarrow p'\colon \tau \in \{\tau\tau\}$$

Finally, acceptance/rejection distributions between s and s' are stable in the now familiar

[8] Frege (1893, §20). Note however that $\equiv$ is a simple predicate, not a defined one as usual.

sense.[9]

$$p\colon E_\Phi, p'\colon \Phi \Rightarrow p\colon \Phi \qquad p\colon E_\Phi, p\colon \Phi \Rightarrow p'\colon \Phi$$

2 Charity

2.1 How universes expand, part I

No doubt these patterns may seem rather recherché. But I take it that there is nothing to stop the Extenders coming to use their new language in this way. And indeed *mutatis mutandis* there would be nothing to stop a sufficiently committed band of initially English-speaking relativists from adopting analogous patterns. The question remains however whether this achieves anything to advance their cause; in particular, whether doing so brings about the expansion of the initial universe that the Extenders set out to achieve. Here we shall give the first half of an argument that it does. On the basis of a substantive assumption about how use determines meaning, we can demonstrate that the following characteristically relativist claim comes out true under the new interpretation.

$R_{p'}$ $\neg\forall x(x \prec mm_p)$

The term mm_p abbreviates the plural term $\langle\lambda_p y.\exists x(x = y)\rangle$ that denotes the members of the initial universe under its intended interpretation. As such, under the intended interpretation of its constituent expressions, $R_{p'}$ says that the initial universe fails to be comprehensive. After pausing in the next section to defend the key assumption, we return to complete the argument in section 3.2.

The key assumption we deploy is a relatively circumscribed principle of Charity, read as a principle about the metaphysics of meaning determination. We assume that the sentences of the Extenders' language are so interpreted that the distributions of acceptance and rejection that members of this community *find* incoherent, really are incoherent—insofar as adopting such a distribution always leads one to accept a falsehood or to reject a truth.

[9] See chapter 4, section 2.2.

More precisely, say that a sequent $\Gamma \Rightarrow \Theta$ is *sound* if there is no assignment σ such that each antecedent sentence is true relative to the perspective denoted by its label under σ and each consequent sentence is false relative to the perspective denoted by its label under σ. Sequent rules are said to be sound if they preserve soundness: the end sequent is sound if the initial sequents are. The assumption we need is simply the following.

Soundness. Ceteris paribus, the sentences of the Extenders' new language are so interpreted as to render the patterns of use they deploy sound.

Of course, if Extenders adopt inconsistent patterns which cannot be soundly interpreted, or other factors involved in meaning determination intervene, their patterns may be unsound; providing however that such things do not occur we assume that meanings are co-ordinated with use in this way.

Assume Bivalence. Notice first that for $M_{p'}$ comprising those objects a such that Ev is true relative to p' when v is assigned to a, this soundness of the following instance of Stability immediately ensures that $M_{p'}$ is at least as inclusive as M_p.

$$p\colon Ev \Rightarrow p'\colon Ev$$

Second it is straightforward to show that $R_{p'}$ comes out true in the new language by running a sequent version of the Russell argument. Write rr to abbreviate $\langle \lambda_p x . x \notin x \rangle$, and r to abbreviate $\{rr\}$. Under their intended interpretations, rr denotes the non-self-membered sets in the initial universe, and r the 'Russell set' with these sets as elements. On the basis of the soundness of the Extenders' patterns of use we may show that Er is not true relative to the initial perspective p, and thus that the target relativist claim, $R_{p'}$, is true relative to the new perspective p'.

The soundness of the Extenders' patterns of use ensures that the end sequents in the following derivations are sound.[10]

[10]Here and below the unlabelled inferences are applications of CUT

$$
\begin{prooftree}
\AxiomC{$p\colon r \in r \Rightarrow p\colon r \in r$}
\RightLabel{L$\neg$}
\UnaryInfC{$p\colon r \in r, p\colon r \notin r \Rightarrow$}
\RightLabel{Lλ}
\UnaryInfC{$p\colon r \in r, p\colon r \prec rr \Rightarrow$}
\AxiomC{$p'\colon r \in r \Rightarrow p\colon r \prec rr$}
\BinaryInfC{$p\colon r \in r, p'\colon r \in r \Rightarrow$}
\AxiomC{$p\colon Er, p\colon r \in r \Rightarrow p'\colon r \in r$}
\BinaryInfC{$p\colon Er, p\colon r \in r \Rightarrow$}
\end{prooftree}
$$

$$
\begin{prooftree}
\AxiomC{$p\colon r \in r \Rightarrow p\colon r \in r$}
\RightLabel{R$\neg$}
\UnaryInfC{$\Rightarrow p\colon r \in r, p\colon r \notin r$}
\AxiomC{$p\colon Er \Rightarrow p\colon Er$}
\RightLabel{Rλ}
\BinaryInfC{$p\colon Er \Rightarrow p\colon r \in r, p\colon r \prec rr$}
\AxiomC{$p\colon r \prec rr \Rightarrow p'\colon r \in r$}
\BinaryInfC{$p\colon Er \Rightarrow p\colon r \in r, p'\colon r \in r$}
\AxiomC{$p\colon Er, p'\colon r \in r \Rightarrow p\colon r \in r$}
\BinaryInfC{$p\colon Er \Rightarrow p\colon r \in r$}
\end{prooftree}
$$

Applying CUT, the left initial sequent of the following derivation is also sound.[11]

$$
\begin{prooftree}
\AxiomC{$p\colon Er \Rightarrow$}
\RightLabel{L $=$}
\UnaryInfC{$p\colon x = r, p\colon Ex \Rightarrow$}
\RightLabel{L$\exists$}
\UnaryInfC{$p\colon \exists x(x = r) \Rightarrow$}
\RightLabel{L$'\lambda_p$}
\UnaryInfC{$p'\colon r \prec mm_p \Rightarrow$}
\AxiomC{$\Rightarrow p\colon Err$}
\AxiomC{$p\colon Err \Rightarrow p'\colon Er$}
\BinaryInfC{$\Rightarrow p'\colon Er$}
\RightLabel{L$\forall\upsilon$}
\BinaryInfC{$p'\colon \forall x(x \prec mm_p) \Rightarrow$}
\RightLabel{R$\neg$}
\UnaryInfC{$\Rightarrow p'\colon \neg\forall x(x \prec mm_p)$}
\end{prooftree}
$$

The conclusion of the end sequent here is the target sentence $R_{p'}$. Consequently, the soundness of these rules ensures that this sentence is true in the Extenders' new language.

2.2 Knowledge maximisation

But why should we expect the Extenders' patterns of use to be sound? I think this question should be answered by defending a more general principle of Charity in interpretation along the lines of the sort defended by Donald Davidson.[12] Rather than treating Charity as a methodological principle governing good interpretation, however, we follow David Lewis (1974) and Timothy Williamson (2004; 2007, ch. 8) in taking some sort of tendency for correctness to be a metasemantic principle, a metaphysical principle about how use determines meanings. This has some initial plausibility. Meanings *do* seem to be determined so as to rule out certain forms of widespread error. Ironic or antiphrastic usages of sentences that become widespread often start to say what they

[11] The rule L$\exists$ may be derived from the rules for $\forall$ and $\neg$ as usual.

[12] See, for instance (1973, 1974a). Davidson's early use of Charity, generalises Quine's (1960, p. 59) limited use of Charity in translation. In 'Radical Interpretation' the principle is formulated in terms of maximising agreement 'in the sense of making [speakers] right, as far as we can tell, as often as possible' (1973, p. 136). But he later comes to accept the refinements offered by Lewis discussed below. See (1974b).

previously merely communicated. Were a linguistic community who started out speaking (the Queen's) English to start using 'bad' as they now use 'good' (either wilfully or erroneously) it seems that it would not be all that long before attribution of good things as 'bad' in that community would cease to be (literally) mistaken.

Evidently, however, widespread and systematic error is possible. This provides a wealth of easy counterexamples to crude principles of Charity that seek simply to maximise correctness.[13]

Correctness Maximisation. Ceteris paribus sentences are so interpreted as to maximise the proportion of acceptance or rejection attitudes towards propositions that are correct.[14]

Contrary to such an injunction to maximise correctness, erroneous computations with hitherto uncalculated-with integers do not run the risk of shifting the meanings of arithmetical expressions like $+$ and $\times$ to render the calculations correct. This remains the case even if the mistake becomes commonplace due, say, to a faulty batch of calculators, or a misprint in a text book. Nor do we have any reason to interpret a technologically primitive linguistic community as speaking exclusively about neutrinos even if the sheer wealth of these particles enables us to always interpret them as being correct by doing so.[15]

But defending Charity does not commit us to defending a crude version of this principle. One option for refinement would be to build an allowance for reasonable error into the '*ceteris paribus*' clause. For example, in his 1974 application of Charity, David Lewis proposes that we make allowance for errors that we might have made in similar circumstances: we should ascribe to the speaker 'those errors which we think we would have made, or should have made, if our evidence and training had been like his.' (1974, p. 336)[16] But in the context of a metaphysical principle about meaning

[13]The formulation in the text adapts the formulation criticised by Williamson (2007). He gives the formulation on page 261.

[14]The correctness of acceptance/rejection is characterised in the obvious way: accepting truths and rejecting falsehoods is correct; rejecting truths or accepting falsehoods is incorrect.

[15]Williamson (2004, pp. 141–4), (2007, pp. 262–6) offers similar and further counterexamples. See also Grandy's (1973) criticism of Quine's (1960) use of Charity applied to translation.

[16]Compare Richard Grandy's (1973, p. 443) 'principle of humanity'. Grandy is primarily

determination this seems overly anthropocentric. As Williamson observes, 'humans are prone to peculiar logical and statistical fallacies: once we recognize a quirky design fault in ourselves, it would be perverse to prefer, on metaphysical principle, interpretations of non-human aliens that attribute the same design fault to them.' (2007, p. 262)

A more promising refinement of the principle due to Williamson (2004; 2007, ch. 8) proposes that we replace correctness maximisation with knowledge maximisation. Following this suggestion, we shall reformulate Charity as follows:[17]

Knowledge Maximisation. Ceteris paribus sentences are so interpreted as to maximise the proportion of acceptance or rejection attitudes towards propositions that constitute knowledge.

So understood, Charity would seem to strike the right balance between permitting new patterns of use to effect semantic change without ruling out the possibility of widespread or systematic error. In the case of 'bad', since speakers will often be in a position to know that the attributed thing is good, Knowledge Maximisation ensures, *ceteris paribus*, that such attributions of good things as 'bad' express truths. In the case of +, since her slips prevent the speaker coming to know the results of her calculations, whether true or not, the mere fact of her acceptance provides no reason for interpreting her to be correct. Similarly since the technologically primitive community are in no position to know about neutrinos, the correctness maximisation of the neutrino-interpretation greatly reduces knowledge.[18]

Let us take the Knowledge Maximisation principle for granted then. How does this relate to Soundness? The answer is that interpreting the basic logical and mathematical inferences that a community makes as sound often tends to maximise knowledge.

Suppose that a linguistic community follows a certain pattern of mental and linguistic behaviour U, with members of this community coming to accept certain proposi-

interested in offering an alternative to Quine's (1960) principle of Charity as a pragmatic constraint on translation.

[17]To adapt Williamson's proposal to the present setting we assume here that rejection attitudes as well as acceptance attitudes can constitute knowledge. The thought here is that even a speaker of a language without negation can come to know—say—that the object in front of her is not red by seeing that it is green and on this basis coming to reject what is expressed by 'that is red'.

[18]This last point is made by Williamson (2007, p. 265–6)

tions and reject others, for a variety of reasons. In particular, we may suppose that some propositions are accepted or rejected on the basis of inference from others in virtue of members of this community conforming to a use pattern of the form $\Gamma \Rightarrow_C \varsigma\colon \phi$. Given U, an interpretation which renders this pattern sound makes such patterns of use a route to inferential knowledge: when speakers know what is expressed by each of the premisses and come to further accept what is expressed by the conclusion on the basis of competently performing the sound inference, they come to know what is expressed by the conclusion. On the other hand, an interpretation which renders the pattern unsound prevents the pattern from being a route to inferential knowledge. When members of C come to accept what is expressed by its conclusion solely through having deduced it from known premisses following what is now an *unsound* inference rule, they do not come to know what is expressed by the conclusion. Even if they chance upon accepting a truth, unsound inference is not a route to knowledge. There is thus a reason to expect Knowledge Maximisation to tend in general towards the soundness of such inferences.

Of course, the increase in inferential knowledge gained by interpreting patterns of inference soundly may have to be traded off against gains in other sorts of knowledge from interpreting them unsoundly. Perhaps, for example, members of a technologically primitive community come to accept what is expressed by the sentences of their language about, say, the prevalence of fish stocks in different parts of a nearby river both directly via perception and indirectly by making dubious inferences based on augury. Knowledge maximisation need not incorrectly require such inferences to be sound, for the inferential knowledge gained under such a deviant interpretation may be outweighed by the perceptual knowledge lost under the actual interpretation.[19]

But in the case of the logical and mathematical patterns of use that our imaginary linguistic community, the Deliberate Extenders, conform to, there seems no reason to

[19]Granted Charity, the assumption that the sentences *are* about the prevalence of fish stocks will only hold good if the potential perceptual knowledge under the one interpretation does outweigh the putative inferential knowledge under the other. If the community tip the balance by coming to rely too heavily on augury, and too little on perception, the sentences in question would come to have truth-conditions unrelated to the distribution of fish. The sentences would in effect come to be about the objects of the augury, or at least what could safely be inferred from the results of such procedures.

think that a similar conflict will arise. Indeed we are at liberty to suppose that their pattern U of mental and linguistic behaviour is such that no such conflict ensues. If need be, we may suppose that aside of accepting logical truths, and rejecting logical falsehoods, all of their acceptance and rejection is based on inference, following the patterns of use set out in section 1.2. In this case, no non-inferential knowledge is gained and inferential knowledge is lost if any of their patterns of use are interpreted as unsound. In this highly idealised case, maximising the proportion of acceptance/rejection attitudes that constitute knowledge leads directly to Soundness. *Ceteris paribus*, Charity as articulated in Knowledge Maximisation vindicates the assumption of Soundness in the account of universe expansion given so far.

2.3 *Ceteris Paribus?*

But is all else equal?

Knowledge Maximisation cannot mandate that speakers' sentences be so interpreted that their patterns of use are sound unless such an interpretation is possible. If, to borrow Prior's famous example,[20] speakers adopt the following patterns of use, then knowledge cannot be maximised by interpreting the speakers' patterns soundly, for these patterns cannot be so interpreted.[21]

$$\varsigma\colon \phi \Rightarrow \varsigma\colon \phi \text{ tonk } \psi \qquad \varsigma\colon \phi \text{ tonk } \psi \Rightarrow \varsigma\colon \psi$$

At the very least, the *ceteris paribus* clause in Soundness must require that a sound interpretation is possible.

So long as we are interested in giving an account of universe expansion from the expansionist's point of view, however, there is an equally straightforward response: unlike the case of 'tonk', the patterns in question clearly *can* be soundly interpreted. From the perspective of a sufficiently liberal metalanguage the expansionist can demonstrate this by constructing a model.[22] The Extenders' patterns of use are sound when the new

[20]Prior (1960).

[21]The empty sequent $\Rightarrow$ may be derived from these rules, and the structural rules, and this is trivially unsound.

[22] Consider the triple $\langle \omega, \omega \times \omega, \{D_n\}_{n \in \omega} \rangle$ which is a frame in the sense of in chapter 5 (section

universe expands the initial universe with all sets of members of the initial universe and the expressions of the Extenders' lexicon are given their intended extensions over the wider universe (with $\lambda_p v\phi$ still interpreted to denote the zero or more members of the *initial* universe that satisfy $\phi(v)$ under the initial interpretation).

Such an argument for soundness is unlikely to impress the absolutist. For, in the crucial case, he contends that the initial universe M_p is absolute. Consequently he denies that the expansionist's model exists. No set, and in particular no set model, has a universal set in its transitive closure.

Moreover, the absolutist may claim that there are simply too few objects to soundly interpret the inter-language version of Basic Law V while according the intended meaning to the identity predicate under the new interpretation. For in order to interpret this pattern in this way, the singular terms $\{xx\}$ and $\{yy\}$ must each denote a distinct member of the new universe whenever xx and yy each denote different members of the initial domain. This requires that there be as many objects in the new universe $M_{p'}$ as there are plural-term denotations over the initial universe M_p. Consequently, applying Cantor's argument, the cardinality of the new universe $M_{p'}$ must exceed that of the purportedly inexpandable initial universe M_p. And this, he concludes, is impossible: no things are more numerous than absolutely everything.[23]

2.1), where D_0 comprises the members of the Extender's initial domain—assumed, recall, to be transitive—and $D_{n+1} = \wp(D_n) \cup D_n$. Consider the model that adds the following extensions: $|p| = 0; |p'| = 1$

$$|v|^\sigma = \sigma(v) \qquad |vv|^\sigma = \sigma(v)$$
$$|\{\tau\tau\}|^\sigma = |\tau\tau|^\sigma \qquad |\langle\lambda_p v.\phi\rangle|^\sigma = \{a \in D_{|p|} \mid |\phi|^{\sigma[v/a]} = T\}$$
$$|{=}|_n = \{\langle a, b\rangle \mid a = b \in D_n\} \qquad |{\equiv}|_n = \{\langle A, B\rangle \mid A = B \subseteq D_n\}$$
$$|{\in}|_n = \{\langle a, b\rangle \mid a \in b \in D_n\} \qquad |{<}|_n = \{\langle a, B\rangle \mid a \in B \subseteq D_n\}$$

Given an assignment σ over the model mapping each singular variable to a member of some D_s and each plural variable to a subset of some D_s, the satisfaction clauses are as in chapter 5 with the following additions.

$|R\tau_1 \ldots \tau\tau_k|_s{}^\sigma = T$ iff $\langle ||_{tau_1}\sigma, \ldots, ||_{\tau\tau_k}\sigma\rangle \in |R|$

$|\forall vv\phi|_s{}^\sigma = T$ iff, for each $A \subseteq D_s$, $|\phi|_s{}^{\sigma[vv/A]} = T$

Then, granted the transitivity of D_0, it is readily verified that the Extenders patterns of use are sound in the model: there is no assignment σ over the model such that the antecedent sentence are true in the model relative to the labelled perspective under σ and the consequent sentences are false in the model relative to the labelled perspective under σ.

[23]Compare Boolos (1997, p. 307) on why the ordinary version of Basic Law V fails. In order

This rejoinder does not interfere with our primary aim of giving an account of universe expansionism acceptable to the relativist. For the relativist does not accept the absolutist's contention that M_p is absolute. It may appear however to frustrate any attempt to turn the relativist's account of universe expansion into an argument for expansionism that has dialectical force against the absolutist. The rejoinder seems to show that the absolutist is under no pressure to accept the possibility of a sound interpretation of the patterns in question.

The dialectical situation however is not so straightforward. Certainly the relativist will agree that the absolutist rejoinder is correct to the extent that the following is true: the patterns of use are sound only if the initial universe fails to be absolute. However, whether the relativist must *presuppose* the initial universe's non-absoluteness turns on a further substantive question in metasemantics. It depends on whether, in the cases in question, subsentential or sentential content has metasemantic priority over the other; whether, in the first instance, Charity operates at a sentential or subsentential level.

Given that the use of expressions determines both the content of whole sentences, and that of subsentential expressions, the question arises whether either category has priority over the other. One possibility is that subsentential expressions have metasemantic priority over sentential expressions. Metasemantics first determines the content of subsentential expressions; and once this has been determined, the content of complex expressions may be recovered in the familiar compositional way. For example, in the case of singular terms we might give a causal story of the sort sketched by Kripke.[24] The meaning of a name is settled prior to the meanings of the whole sentences it occurs in by virtue of there being the right sort of causal connection between tokens of the name and the named object. The other possibility is that whole sentences have metasemantic priority over their subsentential constituents. For example on the sort of account offered by Lewis (1970, 1974), metasemantics operates in the first instance at the level of sentential content. Think of sentential content in an unstructured way, as

to press this objection in good conscience, the absolutist will need to encode the cardinality claims and the proof of Cantor's theorem in plural or higher-order terms along the lines of chapter 2 (section 4).

[24]Kripke (1981, pp. 96–7). See also Devitt (1981).

giving the truth-conditions of sentences; the sort of content that we might model using a set of possible worlds. First of all our use determines the unstructured contents of whole sentences. Given this, we then work backwards to the contents of subsentential expressions. The semantics of subsentential expressions is then so determined as to permit us to compositionally recover the contents of whole sentences that we have already.

Now, if a metasemantic account that thus or otherwise grants priority to subsentential expressions is the right one, then it seems that the absolutist is right to dig in his heels about soundness. For in order to interpret the Extenders' sentences so as to render their patterns of use sound, we must first interpret their subsentential expressions, and—granted absolutism—whatever the relevant metasemantic relation between singular terms and their extensions, singular terms cannot be so related to objects that set terms $\{xx\}$ and $\{yy\}$ denote distinct objects, when xx and yy denote distinct 'pluralities' for the cardinality reasons outlined above.

But in the case of mathematical and logical expressions there is a clear motivation for taking sentential rather than subsentential content to have metasemantic priority. In easy cases, the extensions of our words might plausibly be taken to be concrete objects to which we stand in causal and perceptual relations. Whether or not such theories are ultimately defensible, why a singular term refers to a given concrete object is at least susceptible to a metasemantic story that grants priority to subsentential expressions, relying on a non-semantic link between tokens of the singular term to the referent. On the Kripkean view, for instance, we might hope first to give an account of how by standing in the right sort of perceptual relations to an object when one utters 'I name this such-and-such', the demonstrative 'this' denotes the object in question, and the new name is thereby furnished with the same referent by stipulation; and then to give an account of the sorts of causal relations later uses of the name need bear to the inaugural one in order to refer to the same object. But in the case of logical expressions like $\neg$ it is far from clear how any such account could get off the ground. For according to our best semantic theories, the meaning of $\neg$ (if it is construed as an object at all) is a non-concrete object, such as the truth-function that maps truth to falsehood and vice versa, or the function on sets of worlds which maps each set to its complement, or similar. In the absence of any causal or perceptual relations to such non-concrete objects, we are unable to baptise the denotation of $\neg$ in the manner Kripke suggests. Similar problems arise for mathematical expressions due to our causal and perceptual isolation from their referents.[25]

We face no similar problem if we accord metasemantic priority to whole sentences. For example, assuming Charity calls for soundness, we first determine the truth-conditions of sentences containing $\neg$ so as to render the inferences made with such sentences sound and then on this basis we determine the semantics of the connective itself. Similarly, we first determine the unstructured content of whole sentences containing the set operator $\{\cdot\}$, so as to render the inter-perspective version of Basic Law V sound and then work backwards to the semantics of $\{\cdot\}$ itself. In this way, we are able to determine the semantic values of logical and mathematical expressions without needing to stand in causal or perceptual relations to them.

Now, given the metasemantic priority of whole sentences, the availability of a sound interpretation of the Extenders' patterns of use need not presuppose the existence of objects outside their initial universe. In order to interpret these patterns soundly all we need is to assign suitable truth-conditions to each sentence; a sound interpretation will be achieved by the right sort of distribution of unstructured propositions across the sentences of the language. But the availability of the unstructured contents required for such a distribution does not beg the question against absolutism. It does not presuppose that the initial universe is expandable.[26]

[25]This can be seen as a metasemantic analogue of the epistemological difficulty set out in Benacerraf's famous (1973) paper in reconciling the abstract nature of mathematical objects with causal theories of knowledge.

[26]The argument for universe expansion presented so far has some affinity to arguments put forward by Eli Hirsch in the context of metaontology. Hirsch contends, for example, that one can come to speak a language that quantifies over mereological sums outside the universe of the initial language by specifying truth-conditions for sentences of the new language with sentences of the old language. (Hirsch 2002, p. 54). Unlike the present account, however, Hirsch does not seek to give a systematic account of the relation between the old and new language: rather he thinks that one can learn to use the new language ostensively through a few representative examples (p. 58). Hirsch also makes use of charity, although in a rather different way to here, as a means to link use to meaning. See Hirsch (2005).

3 Content Determination

The appeal to the metasemantic priority of sentential content, however, raises a further worry. For consider the account of universe expansion given so far. Soundness requires that the contents of whole sentences be determined so as to render patterns of use sound; and when applied to the patterns of inference adopted by the Extenders, this ensures that the relevant instance of the schematic statement of relativism

$R_{p'}$ $\neg\forall x(x \prec mm_p)$

comes out true under the new interpretation. This, however, is not enough to show that the Extenders have succeeded in expanding the initial universe. For, as yet, nothing has been said about the interpretation of the Extenders' *sub*sentential expressions. But whether or not universe expansion occurs is a matter of subsentential interpretation: it is a question of whether or not new objects, objects outside the initial universe, come to be the extensions of singular terms, or come to be in the extensions of predicates, and so on. The truth of $R_{p'}$ on its own does not show that this has been achieved. Of course, one way for R'_p to be true is for the subsentential expressions in this sentence to receive their intended interpretations over a strictly wider universe. But there are many other ways that the subsentential expressions could be re-interpreted in order to achieve this result. For example, all of the expressions bar $\neg$ could receive their intended interpretation over the *initial* universe; and $\neg$ could be reinterpreted to have the intended interpretation of $\neg\neg$, so that under the resulting unintended interpretation, $\neg\phi$ is true just in case ϕ is true. Rather than stating that there is a new object outside the initial universe, under the deviant interpretation, $R_{p'}$ reaffirms the comprehensiveness of this universe.

In order for the foregoing account to be an idealised explanation of universe expansion, the end result of adopting the relevant patterns of use must be universe expansion. Unless we can show that the sentential interpretation mandated by Soundness secures the intended semantics of the expressions in $R_{p'}$, the account of universe expansion fares little better than an amateur ontological argument that proceeds by stipulating that 'God exists' is to express some trivial truth. And this may initially seem hopeless: granted that Soundness operates at the level of sentential content, there is no unique backwards route from the content of a whole sentence to the content of its subsentential expressions.

3.1 Logical content determination

We should not however be unduly pessimistic. While both these points are true, the worry overlooks the fact that Soundness constrains the sentential content not just of $R_{p'}$ but of the other sentences in the language containing the subsentential expressions in this sentence. In the case of $\neg$ for example, a sound interpretation of the Extenders' patterns of use will render sound the following immediate consequences of the structural rules and the operational rules for this connective.

$$p'\!:\phi, p'\!:\neg\phi \Rightarrow \qquad\qquad \Rightarrow p'\!:\phi, p'\!:\neg\phi$$

And the soundness of these sequents rules out the above unintended interpretation of $\neg$ as expressing double negation. In fact, assuming a bivalent interpretation, the soundness of the two sequents suffices to secure the intended interpretation for this connective. Soundness of the left sequent ensures that not both ϕ and $\neg\phi$ are true under p'; soundness of the right sequent ensures that not both ϕ and $\neg\phi$ are false under p'. Given that each sentence is true or false under p', it follows that $\neg\phi$ is true under p' if and only if ϕ is not true under p'.

This result can be generalised to the other logical expressions in the Extenders' lexicon (including λ_p-terms). Vann McGee (2000, 2006) has shown that given a minimal assumption—and under an absolutist-friendly, plural construal of 'model'—that if the classical natural deduction inference rules, read open-endedly, preserve truth under every model, we can establish the familiar Tarskian truth-conditions for the connectives, identity predicate and first and second-order quantifiers. Now the relativist will not accept that the classical rules preserve truth when understood open-endedly. For as in chapter 4 (section 2.1), Universal Specification fails to preserve truth when the term substituted for the bound variable denotes an object outside the domain of the quantifier, as the relativist contends it always may. However, a similar result may be established in the present setting. An analogue of McGee's result can be proved when we replace the natural deduction rules for classical second-order logic with the sequent rules characteristic of free plural logic FPFO displayed in table 6.3, above, and truth-preservation under every model with (material) soundness.

The result makes surprisingly few assumptions about the background semantics. We use only the following assumption about truth under an assignment (which is analogous to the assumption employed by McGee).

The irrelevance of assignments to non-free variables When assignments σ and ρ agree on all the free variables in a sentence ϕ, ϕ is true under p' and σ if and only if ϕ is true under p' and ρ.

As in the proof of McGee's theorem, we reverse the familiar order of explanation. Rather than showing that the intended semantics renders the rules sound, we show that the soundness of the rules for p'—together with the soundness of the substitutivity rules for the identity predicate in English—pins down the intended semantics. We have seen the case for $\neg$ already. Assuming bivalence, the soundness of the patterns of L$\neg$ and R$\neg$ and the structural rules ensure that it receives its intended semantics. In fact we may drop the assumption of Bivalence. For the soundness of the structural rules rules out truth-value gaps and gluts. Given this, the p'-instances of the characteristic rules for each logical expression ensure that it receives the intended satisfaction clause relative to p'. More precisely, we have the following theorem.

Content Determination Theorem. Assume the irrelevance of free variables that do not occur free and that the substitutivity rules for English singular and plural identity predicates are sound. Suppose that instances of the patterns of use: REP, DIL, CUT, L$\neg$, R$\neg$, L$\wedge$, R$\wedge$, L $=$, R $=$, L $\equiv$, R $\equiv$ L$\forall v$, R$\forall v$, L$\forall vv$ and R$\forall vv$, displayed in table 6.3 (on page 197 above) and L$'\lambda_p$ and R$'\lambda_p$ (on page 198) are sound when ς is replaced with p'. Then the satisfaction clauses for the Extenders expressions displayed in table 6.4 hold for all assignments σ.[27]

[27] Notice that in the case of λ_p-terms and the member-'plurality' predicate, the intended truth-conditions for sentence of the form $u \prec \langle \lambda_p v.\phi \rangle$ do not suffice to simultaneously pin down the

Table 6.4: Intended truth conditions for logical expression relative to p'

Let $\mathrm{Den}_\sigma(vv)$ abbreviate 'the (zero or more) objects σ maps vv to'; $\mathrm{Den}_\sigma(\langle\lambda_p v.\phi\rangle)$ abbreviate 'the zero or more members a of M_p such that $\phi[u/v]$ is true relative to p under $\sigma[a/u]$'.

$u = v$ is true relative to p' under σ iff $\sigma(u)$ is identical to $\sigma(v)$, which is a member of $M_{p'}$

$\tau\tau_1 \equiv \tau\tau_2$ is true relative to p' under σ iff $\mathrm{Den}_\sigma(\tau\tau_1)$ are identical to $\mathrm{Den}_\sigma(\tau\tau_2)$, which are members of $M_{p'}$

$u \prec \langle\lambda_p v.\phi\rangle$ is true relative to p' under σ iff $\sigma(u)$ is one of $\mathrm{Den}_\sigma(\langle\lambda_p v.\phi\rangle)$.

$\neg\phi$ is true relative to p' under σ iff ϕ is not true relative to p' under σ.

$\phi_1 \wedge \phi_2$ is true relative to p' under σ iff ϕ_1 is true and ϕ_2 is true relative to p' under σ.

$\forall v\phi$ is true relative to p' under σ iff, for every a in $M_{p'}$, ϕ is true relative to p' under $\sigma[v/a]$.

$\forall vv\phi$ is true relative to p' under σ iff, for all zero or more aa in $M_{p'}$, ϕ is true relative to p' under $\sigma[vv/aa]$.

Note that the proof makes essential use of open-endedness, exploiting the inter-perspective nature of the substitutivity rules for $=$ and $\equiv$ and allowing side-formulas to be labelled with perspective terms other than p'. The proof is outlined in the appendix (p. 223ff.)

What of the auxiliary assumption? If we are reinterpreting the lexicon in the first instance by settling the content of whole sentences, we can in principle stipulate interpretations that violate the irrelevance of non-free variables. For example, if the Extenders were to adopt patterns of use like the following then soundness would require that the irrelevance of non-free variables be violated.

$$p\colon Fx \wedge x = y \Rightarrow p'\colon Fx \qquad p'\colon Fx \Rightarrow p\colon Fx \wedge x = y$$

The truth of Fx relative to p' would depend not only on the assignment of x but also on the assignment of y.

Crucially, however, members of our imaginary linguistic community do no such thing. All of the patterns of use for atomic sentences co-ordinate the acceptance or rejection of sentences as interpreted relative to the new perspective with that of sentences containing the same free variables relative to the initial perspective. For example, the

intended extensions of both $\prec$ and $\langle\lambda_p v.\phi\rangle$. A deviant interpretation of $\prec$ can be compensated for by a deviant interpretation of $\langle\lambda_p v.\phi\rangle$. However for present purposes the intended truth-conditions for member-'plurality' statements of the canonical form will suffice.

soundness of the patterns of use adopted for set-set identities ensure that the truth-conditions of $\{xx\} = \{yy\}$ in the new language are co-ordinated with those of $xx \equiv yy$ in the initial language. Since the truth-value of the latter is invariant under changes to the assignment of variables other the xx and yy, the same is true of the former. Granted the intended interpretation of the logical expressions, it is then readily verified that this property is inherited by more complex expressions. This goes to show that sound interpretations of the Extenders' patterns need not violate the irrelevance of assignments to non-free variables. Consequently, we can force such an interpretation by supposing—with considerable plausibility—that, when σ and ρ agree on the variables in ϕ, the Extenders find it incoherent to accept what is expressed by an ϕ under p' and σ and to reject what is expressed by ϕ under p' and ρ. The soundness of this pattern—in the obvious extended sense[28]—ensures that the auxiliary assumption is met.

3.2 How universes expand, part II

The Content Determination Theorem provides the means to complete the argument for universe expansion. Recall the first part of the argument for universe expansion from section 2.1 above. The soundness of the Extenders' patterns of use ensures first that $M_{p'}$ is at least as inclusive as M_p and second that

$R_{p'}$ $\neg\forall x(x \prec mm_p)$

is true relative to the new perspective p'. The Content Determination Theorem ensures that this sentence receives its intended truth-condition. Working through the truth-conditions in table 6.4, $R_{p'}$ is true relative to p' just in case it is not the case that every member a of the new universe $M_{p'}$ is one of the zero or more members of M_p such that $\exists x(x = y)$ is true relative to p when y is assigned to a. Granted that the initial language receives its intended interpretation over M_p, as above, this holds just in case not every member of the new universe is a member of the initial universe. Putting the two parts together, the new universe properly expands the initial one. Given Soundness, the Extenders' new patterns of use cause their universe to expand.

[28]Namely, the assumption that there are no assignments σ and ρ related in the specified way, such that Φ is true relative to p' and σ and Φ is false relative to p' and ρ.

This opens the way for a charity-based metasemantics to provide a straightforward means for universe expansion to occur. Expanding the universe is simply a matter of coming to use language in the right way. If a linguistic community adopts a pattern of use that can only be charitably interpreted over a wider universe—providing all else is equal—their lexicon is interpreted over the wider universe; and their unrestricted quantifiers thereby come to quantify over new objects.

Charity also provides the bridge between use and meaning for the interpretational expansionist to meet the side-condition problem in the manner outlined in chapter 3, section 2.3. In order to formulate relativism using the schema template

$(R_\varsigma)\ \varsigma' \colon \neg\forall x(x \prec mm_\varsigma)$

the relativist needs to be able to frame the side-condition that ς' denote a suitable meta-perspective for the perspective denoted by ς in a way that is both semantically adequate, ensuring that instances of this sentence receive their intended truth-conditions, and conceptually licit, deploying only expressive resources her view permits.

But granted that Charity ensures Soundness, use provides a conceptually licit handle on semantic adequacy. Whether one perspective is a meta-perspective for another can be characterised in terms of the inter-perspective inferential relations between the two. In the idealised case of the Extenders, a sufficient condition for p' to be a meta-perspective for p is that members of the community conform to the patterns of use set out in section 1.2. This is conceptually licit: describing such patterns of use does not call for absolutely general quantification, but only circumscribed generality about the the linguistic and mental behaviour of a particular linguistic community. But once we have Soundness via Charity, the condition is also semantically adequate. By the Content Determination Theorem, the expressions in $R_{p'}$ receive their intended semantics over a universe that is more inclusive than M_p.

4 Limits

Let us take stock. Charity, understood as Knowledge Maximisation, is not susceptible to the problems that face rival formulations of this principle. Given this substantial metasemantic assumption, expanding the universe—like effecting any other sort of shift in meaning—is simply a matter of coming to adopt the right patterns of use. In particular, in the case of an idealised linguistic community we saw how the soundness of their patterns of use ensures that the logical component of the lexicon receives its intended interpretation over a universe that is strictly more inclusive than the initial one. But while this provides the relativist with an account of universe expansion, she might more ambitiously attempt to use it as an argument in favour of the same. For provided that Charity operates in the first instance at the level of the semantic content of whole sentences, the availability of a sound interpretation of the Extenders' patterns of use need not presuppose the availability of a wider universe.

There is however an important caveat that has yet to be made concerning this second, more ambitious use of the story about universe expansion. It does not seem to be available to all expansionist relativists. This style of argument against the absolutist's claim that some singular quantifiers $\forall v$ achieve absolute generality, can also be pressed against the claim that 'modalised' quantifiers, strings of the form $\Box\forall v$, achieve absolute generality, which is characteristic of pluralist relativists. Consequently, this sort of account of universe expansion pushes the relativist towards a thorough-going version of her view according to which, not only quantifiers but also modalised quantifiers fail to range over an inexpandable universe.

To see this, let us return to the idealised linguistic community. Suppose that following their apparent success in expanding the initial universe, the Extenders have iterated this expansion, and have come to be committed to schematic versions of their earlier patterns of use.

$$\varsigma'\colon \{\tau\tau_1\} = \{\tau\tau_2\} \Rightarrow \varsigma\colon \tau\tau_1 \equiv \tau\tau_2 \qquad \varsigma\colon \tau\tau_1 \equiv \tau\tau_2 \Rightarrow \varsigma'\colon \{\tau\tau_1\} = \{\tau\tau_2\}$$

$$\varsigma'\colon \tau \in \{\tau\tau\} \Rightarrow \varsigma\colon \tau \prec \tau\tau \qquad \varsigma\colon \tau \prec \tau\tau \Rightarrow \varsigma'\colon \tau \in \{\tau\tau\}$$

$$\varsigma\colon E_\Phi, \varsigma'\colon \Phi \Rightarrow \varsigma\colon \Phi \qquad \varsigma\colon E_\Phi, \varsigma\colon \Phi \Rightarrow \varsigma'\colon \Phi$$

Suppose further that they have come to enrich their lexicon with an inter-perspective modal operator: $\Box$. To state their patterns of use for the modal operator, we need to slightly generalise the notation. The labelled sequent $R\varsigma\varrho, \Gamma \Rightarrow_C \Theta$ may be glossed as saying that when the perspective labelled by ϱ is accessible from the perspective labelled by ς, members of community C find it incoherent to accept the antecedent propositions and reject the consequent ones. The Extenders have come to adopt the follow patterns of use for $\Box$.

$$\frac{\varsigma\colon \phi, \Gamma \Rightarrow \Theta}{R\varsigma\varrho, \varrho\colon \Box\phi, \Gamma \Rightarrow \Theta}\,\text{L}\Box \qquad \frac{R\varrho\upsilon, \Gamma \Rightarrow \Theta, \upsilon\colon \phi}{\Gamma \Rightarrow \Theta, \varrho\colon \Box\phi}\,\text{R}\Box$$

The right-hand rule is subject to the side condition that υ be replaced with a perspective variable that does not occur in the end-sequent. The availability of such a modal operator, however, leads to a fresh controversy amongst the imaginary community. There is a fresh schism between those who agree with pluralist relativists in thinking that even though the domain of any unrestricted quantifier $\forall v$ relative to any perspective r fails to be inexpandable, one can nonetheless achieve absolute generality relative to this perspective by using a modalised quantifier $\Box\forall v$, and those who agree with thorough-going relativists that neither quantifiers nor modalised quantifiers achieve such generality. Encouraged by their previous success, the thorough-going sub-community decides to attempt to expand the range of modalised quantifiers.

To do this they first adopt patterns of use that co-ordinate acceptance and rejection relative to the perspectives ranged over by perspective variables υ, to a new perspective $p^\star$. Let us suppose that they enrich their lexicon with a plural-term forming operator $\lambda_{p^\star}$ which given a variable v and a sentence ϕ yields a plural term $\langle \lambda_{p^\star} v.\phi \rangle$. Moreover they come to accept the $p^\star$-instances of the sequent rules set out in table 6.3 (on page 197) including in particular the rules for the new $\lambda_{p^\star}$ terms. Given $p^\star$, they iterate the same patterns for set terms before, continuing to conform to the instances of the patterns for $p^\star$, $p^{\star\prime}$, $p^{\star\prime\prime}$ and so on. They also adopt analogous pattern to before governing $\lambda_{p^\star}$-terms at $p^{\star\prime}$

$$\frac{p^\star\colon E\tau, p^\star\colon \phi[\tau/v], \Gamma \Rightarrow \Theta}{p^{\star\prime}\colon \tau \prec \langle \lambda_p v.\phi \rangle, \Gamma \Rightarrow \Theta}\,\text{L}'\lambda_p \qquad \frac{\Gamma \Rightarrow \Theta, p^\star\colon \phi[\tau/v] \qquad \Gamma \Rightarrow \Theta, p^\star\colon E\tau}{\Gamma \Rightarrow \Theta, p^{\star\prime}\colon \tau \prec \langle \lambda_p v.\phi \rangle}\,\text{R}'\lambda_p$$

Finally they extend the rules for $\Box$ to the new perspective terms and adopt the following version of stability for the new perspective:

$$Rpr, r'{:}\, E_\Phi, p^\star{:}\, \Phi \Rightarrow r'{:}\, \Phi \qquad Rpr, r{:}\, E_\Phi, r{:}\, \Phi \Rightarrow p^\star{:}\, \Phi$$

$$p^\star{:}\, E_\Phi, p^\star{:}\, \Phi \Rightarrow p^{\star\prime}{:}\, \Phi \qquad p^\star{:}\, E_\Phi, p^{\star\prime}{:}\, \Phi \Rightarrow p^\star{:}\, \Phi$$

The soundness of these patterns calls for a plural term denotation outside the range of the modalised quantifier.[29] The argument is much the same as before. Let $rr^\star$ abbreviate $\langle \lambda_{p^\star} x.x \notin x\rangle$ and $r^\star$ abbreviate $\{rr^\star\}$. Notice first that since their patterns of use between $p^\star$ and $p^{\star\prime}$ exactly mirror the earlier ones between p and p', replacing the latter with the former in the first two derivation in section 2.1 yields a derivation of $p^\star{:}\, Er \Rightarrow$ by CUT. We then reason as follows.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{r{:}\, Err^\star \Rightarrow r'{:}\, Er^\star \qquad Rps, s'{:}\, Er^\star \Rightarrow p^\star{:}\, Er^\star}{Rps, s{:}\, Err^\star \Rightarrow p^\star{:}\, Er^\star} \qquad p^\star{:}\, Er^\star \Rightarrow}{Rps, s{:}\, Err^\star \Rightarrow}\;\text{L}\equiv}{Rps, s{:}\, yy \equiv rr^\star, s{:}\, Eyy \Rightarrow}\;\text{R}\neg}{Rps, s{:}\, Eyy \Rightarrow s{:}\, yy \not\equiv rr^\star}\;\text{R}\forall\upsilon}{Rps, \Rightarrow s{:}\, \forall yy(yy \not\equiv rr^\star)}\;\text{R}\Box}{\Rightarrow p{:}\, \Box\forall yy(yy \not\equiv rr^\star)}$$

The soundness of the penultimate sequent ensures that $\forall yy(yy \not\equiv rr^\star)$ is true relative to any perspective that is accessible from p. But, as before, granted the soundness of the patterns of use for logical expressions, the Content Determination Theorem ensures that, relative to each perspective q, the expressions in this sentence receive their intended interpretations over the universe M_q of that perspective. Consequently, the objects denoted by rr are not in the universe of any perspective accessible from p. It follows that these objects falls outside the range of the modalised quantifiers $\Box\forall yy$. And indeed the soundness of the final sequent tells us that $\Box\forall yy(yy \not\equiv rr^\star)$ is true relative to p.

The pluralist consequently finds herself in a similar position to the absolutist. Given the soundness of certain straightforward patterns of use, her characteristic claim about the modalised plural quantifiers is false. We can come to a perspective whose

[29] To allow for patterns of use with sentences labelled by perspective variables, we need to take assignments to map these to perspectives in addition to assigning values to singular and plural variables. The definitions of soundness is then as before, where the perspective denoted by a perspective variable υ under σ is $\sigma(\upsilon)$.

universe contains some things that the modalised quantifier $\Box\forall vv$ does not range over relative to the initial perspectives. As before, the thorough-going relativist can demonstrate the possibility of a sound interpretation from a sufficiently liberal metalanguage with a model.[30] The through-going Extenders' patterns of use are sound when the new universe $M_{p^\star}$ comprises all members of any of the initial universes M_r, and the Extenders patterns receive their intended extensions over each universe.

Of course, the pluralist relativist will deny that there is any such model. No set contains a set comprising every member of every initial universe in its transitive closure. Indeed no things comprise every member of every initial universe. But as before it is not obvious that the thorough-going expansionist must presuppose this in order to make a case for the possibility of a sound interpretation. If Soundness operates at the level of sentential contents, obtaining a sound interpretation is simply a matter of distributing unstructured sentential contents across the sentences of the language in the right way. The case against pluralism from the charity-based argument exactly parallels that against absolutism. It seems therefore that this argument cannot be a dialectically effective tool for the pluralist to wield against the absolutist. *If* the argument is effective in showing that the relativist is right to reject the absolutist's claim that unrestricted quantifiers achieve absolute generality, it is equally effective in showing that the thorough-going relativist is right to reject the pluralist's claim that some modalised quantifiers achieve absolute generality. The charity-based metasemantics pushes us towards relativism about modalised quantifiers just as much as relativism about quantifiers.

[30]Consider the frame $\langle \omega + \omega, R, \{D_\alpha\}_{\alpha\in\omega+\omega}\rangle$, where, D_0 comprises the members of M_p, as before, $D_{\alpha+1} = \wp(D_\alpha) \cup D_\alpha$; $D_\omega = \bigcup_n D_n$;

$$R = \{\langle n, m\rangle : n, m < \omega\} \cup \{\langle \alpha, \beta\rangle \mid \omega \leqslant \alpha, \beta < \omega + \omega\}$$

Let the extensions of the Extenders' expressions and the satisfaction conditions be the same as in n. 22, extended to the new perspectives, with the following additions. $|p^\star|^\sigma = \omega$; $|r|^\sigma = \sigma(r)$ $|\langle \lambda_{p^\star} v.\phi\rangle|^\sigma = \{a \in D_{|p^\star|} \mid |\phi|_{p^\star}{}^{\sigma[v/a]} = T\}$ and $|\Box\phi|_s{}^\sigma = T$ iff, for each t such that $\langle s, t\rangle \in R$, $|\phi|_t{}^\sigma = T$. It is then easily seen that the Extenders' patterns outlined in this section together with those cultivated by the thorough-going group for $p^\star$ are sound in the resulting model.

Concluding Remarks

Let us distinguish between three ways we might group objects together. Most tenuously, we can use a condition ϕ to distinguish each object that satisfies the condition from each object that does not. More demandingly, we can collect the objects together, by using a plural term such as $\langle\lambda_p v.\phi\rangle$ or vv to refer to them severally. Third, and most substantially, we can bring their collection into the reach of singular reference, via a set that has them as elements, which may be denoted by singular terms such as those of the form $\{\langle\lambda_p v.\phi\rangle\}$, $\{vv\}$ or v.

In contending that we can achieve unrestricted singular quantification and full unrestricted plural quantification over an inexpandable universe, the absolutist distinguishes the first two from the third. Whenever we can draw a lasso in logical space with a condition ϕ, the objects thus circumscribed can be collected together severally as the objects denoted by $\langle\lambda v.\phi\rangle$; but the objects thus collected together need not have a collection. Indeed in certain crucial cases, like the things that comprise everything or the sets that comprise every non-self-membered set, the absolutist contends that there cannot be such a collection. If the more ambitious aim of this chapter is successful, the Charity-based argument shows this mismatch between the second and third ways of collecting to be unstable. Whenever we can refer to some things severally relative to one perspective, we can refer to their set relative to another. The ability to code some objects as the denotation of a plural term, permits us to develop patterns of use in which the same objects are encoded as the denotation of a set term. The sound interpretation of such patterns leads to the expansion of the universe with a further rank of sets, permitting us to come to refer to and quantify over an object, indeed many objects, that falls outside the range of the purportedly absolutely general quantifier $\forall v$.

Granted that this procedure can be iterated indefinitely, we must give up on the absolutist's contention that some universe is ultimate. There is an endless sequence of languages with wider and wider universes. Accordingly, absolute generality cannot be attained by quantification over a universe. Still, a moderate-seeming sort of relativist might seek to save absolute generality with further expressive resources. Granted a modal operator encompassing every language in the endless sequence of languages, she claims that absolute generality is attained with modalised quantifiers: $\Box\forall v$ ranges over any object that is an urelement or a set in any universe; $\Box\forall vv$ ranges over whatever objects can be collected together severally in any universe. On such a view, there is no longer a mismatch between the the second and third means of collecting: whatever objects are denoted by a plural variable xx are likewise encoded by the denotation of the corresponding set term $\{xx\}$. But in collapsing this divide we open up another. In enriching the language with the modal operators, we can formulate further conditions ψ containing these operators. And insofar as the modalisation of Plural Comprehension fails the first method of collecting no longer lines up with the second. But once again this proves to be unstable in the light of Charity. We can adopt patterns of use whose soundness forces members of the universe of any language within the scope of $\Box$ into a single universe. In effect this provides a means of taking the limit of the infinite sequence of languages encompassed by $\Box$. But in thus merging the universes of these languages, we gain the ability to refer to some things plurally that are not jointly available in any of the initial universes, and are thus outside the range of the purportedly absolutely general quantifier $\Box\forall vv$.

It is in this way that the universe expansion argument pushes us towards a thorough-going version of relativism. Given unrestricted singular and plural quantification over a universe, we can adopt patterns of use whose soundness requires a strictly wider universe. Given a modal operator ranging over an unbounded sequence of universes, we can adopt patterns of use whose soundness requires a universe as inclusive as each. But at no stage can we complete the process. Further expressive resources, in the form of quantification over wider universes, or modal operators encompassing longer sequences of languages, give us the expressive power to adopt patterns of use whose sound interpretation demands yet wider universes, or yet further languages, outside their reach.

5 Appendix: Proof of the Content Determination Theorem

Observe first that the structural rules rule out truth-value gaps and gluts.

> *No gluts.* The soundness of REP immediately ensures that there is no assignment under which ϕ is both true and false relative to p'
>
> *No gaps.* The open-ended soundness of CUT ensures that ϕ is either true or false relative to p' under each assignment. Let σ be given and suppose for *reductio* that ϕ is neither true nor false relative to p' under σ. Consider sentences of the form $A^w_\sigma u$, true relative to perspective q under ρ iff $\rho(w) = \sigma(u)$. (For example the English sentence 'w is is identical to $\sigma(u)$', as interpreted in English) Then both the following sequents are sound.
>
> $$\{q{:}\, A^v_\sigma v \mid v \text{ a variable}\}, p'{:}\, \phi \Rightarrow \qquad \{q{:}\, A^v_\sigma v \mid v \text{ a variable}\} \Rightarrow p'{:}\, \phi$$
>
> This is because when ρ satisfies $A^x_\sigma x, A^y_\sigma y, \ldots$, relative to q, ρ is identical to σ and thus the remaining antecedent sentence of the left sequent is not true relative to p' under ρ and the consequent sentence of the right sequent is not false relative to p' under ρ, *ex hypothesi.* The soundness of CUT then ensures that the following sequent is sound.
>
> $$\{q{:}\, A^v_\sigma v \mid v \text{ a variable}\} \Rightarrow$$
>
> But this sequent is not sound. Each antecedent sentence is true relative to q under σ

We then check each satisfaction clause in turn.

> The case for $\neg$ is as in the text (via no gaps).
>
> $\wedge$. The soundness of L$\wedge$ and R$\wedge$and the structural rules ensures that the following sequents are sound
>
> $$p'{:}\, \phi_1, p'{:}\, \phi_2 \Rightarrow p'{:}\, \phi_1 \wedge \phi_2 \qquad p'{:}\, \phi_1 \wedge \phi_2 \Rightarrow p'{:}\, \phi_1 \qquad p'{:}\, \phi_1 \wedge \phi_2 \Rightarrow p'{:}\, \phi_2$$

The satisfaction clause for $\wedge$, for all assignments follows from their soundness (by no gaps).

$=$. By the soundness of the rules for $=$ in English and L $=$ and R $=$ and the structural rules, the following sequents are sound.

$$p'\colon u = v, \textit{English}\colon u \text{ is identical to } u \Rightarrow \textit{English}\colon u \text{ is identical to } v$$
$$p'\colon u = v \Rightarrow p'\colon Ev$$
$$\textit{English}\colon u \text{ is identical to } v, p'\colon u = u \Rightarrow p'\colon u = v$$

Suppose $u = v$ is true relative to p' under σ. Since $\sigma(u) = \sigma(u)$, the antecedent sentences in the first sequent are both true relative to the labelled perspectives under σ, so (by no gaps), 'u is identical to v' is true in English under σ, so $\sigma(u) = \sigma(v)$. Moreover the soundness of the second sequent ensures $\sigma(v)$ is one of $M_{p'}$. Conversely, if $\sigma(u) = \sigma(v)$ and $\sigma(v)$ is one of $M_{p'}$, then the antecedent sentences of the third sequent above are true relative to their labelled perspectives under σ, so (by no gaps) $u = v$ is true relative to p' under σ.

Similarly for $\equiv$.

λ_p. The soundness of L$'\lambda_p$ and R$'\lambda_p$ and the structural rules ensures that the following sequents are sound.

$$p'\colon u \prec \langle \lambda_p v.\phi \rangle \Rightarrow p\colon Eu$$
$$p'\colon u \prec \langle \lambda_p v.\phi \rangle \Rightarrow p\colon \phi[u/v]$$
$$p\colon Ev, p\colon \phi[u/v] \Rightarrow p'\colon u \prec \langle \lambda_p v.\phi \rangle$$

Suppose $u \prec \langle \lambda_p v.\phi \rangle$ is true relative to p' under σ. Then the soundness of the first two sequents ensures that $\sigma(u)$ is a member of M_p and $\phi[u/v]$ is true relative to p under σ. Consequently, $\sigma(u)$ is one of the zero or more members a of M_p such that $\phi[u/v]$ is true relative to p under $\sigma[u/a]$. The converse follows by the soundness of the third sequent.

$\forall v$. Let σ be given. Suppose $\forall v\phi$ is true relative to p' under σ. Let a be a member of $M_{p'}$. Then $\forall v\phi$ is true relative to p' under $\sigma[a/v]$ by the irrelevance of non-free variables and Ev is true relative to p' under $\sigma[a/v]$ by the definition of $M_{p'}$.

The soundness of L$\forall v$ and the structural rules ensures that the following sequent is sound

$$p' : \forall v\phi, p' : Ev \Rightarrow p' : \phi$$

Since the antecedent sentences are true relative to p' under $\sigma[a/v]$, ϕ is true relative to p' under $\sigma[a/v]$ (by no gaps).

Conversely suppose ϕ is true relative to p' under $\sigma[a/v]$ for each a in M_p. Then the following sequent is sound, with $A^v_\sigma v$ and q as above.

$$\{q : A^u_\sigma u \mid u \text{ distinct from } v\}, p' : Ev \Rightarrow p' : \phi$$

This is because when ρ satisfies each sentence labelled in $\{q : A^u_\sigma u \mid u \text{ distinct from } v\}$ relative to q, ρ is identical to $\sigma[a/v]$, where $a = \rho(v)$. When moreover Ev is true relative to p' under ρ, a is in $M_{p'}$, therefore ϕ is true under $\sigma[a/v]$, *ex hypothesi*, and thus not false (by no gluts). Consequently, by the soundness of R$\forall v$, the following sequent is also sound.

$$\{q : A^u_\sigma u \mid u \text{ distinct from } v\} \Rightarrow p' : \forall v\phi$$

Since σ satisfies each of the antecedent sentences relative to q, $\forall v\phi$ is not false relative to p' under σ, and thus true (no gaps).

$\forall vv$. Similarly.

Bibliography

Ajdukiewicz, K.: 1935, Die syntaktische Konnexität, *Studia Philosophica* pp. 1–27. Translation in English as 'Syntactic Connexion' in S. McCall (ed) *Polish Logic*, Clarendon Press, Oxford, 1967.

Bach, K.: 2000, Quantification, qualification and context: A reply to Stanley and Szabo, *Mind & Language* **15**(2-3), 262–283.

Bar-Hillel, Y.: 1964, *Language and Information*, Addison-Wesley Publishing Company, Reading, MA.

Barwise, J. and Cooper, R.: 1981, Generalized quantifiers and natural language, *Linguistics and philosophy* **4**(2), 159–219.

Benacerraf, P.: 1973, Mathematical truth, *The Journal of Philosophy* **70**(19), 661–679.

Benacerraf, P. and Putnam, H. (eds): 1983, *Philosophy of Mathematics: Selected Readings*, 2nd edn, Cambridge University Press, Cambridge.

Bencivenga, E.: 1986, Free logics. In D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic*, Kluwer, Dordrecht.

Boolos, G.: 1971, The iterative conception of set, *The Journal of Philosophy* **68**(8), 215–231. Reprinted in Boolos (1998). Page references to reprint.

Boolos, G.: 1984, To be is to be a value of a variable (or to be some values of some variables), *The Journal of Philosophy* **81**(8), 430–449. Reprinted in Boolos (1998).

Boolos, G.: 1985, Nominalist platonism, *The Philosophical Review* **94**(3), 327–344. Reprinted in Boolos (1998).

Boolos, G.: 1989, Iteration again, *Philosophical Topics* **17**, 5–21. Reprinted in Boolos (1998). Page references to reprint.

Boolos, G.: 1993, Whence the contradiction?, *Aristotelian Society Supplementary Volume* **67**, 213–233. Reprinted in Boolos (1998). Page references to reprint.

Boolos, G.: 1995, *The Logic of Provability*, Cambridge University Press, Cambridge.

Boolos, G.: 1997, Is Hume's principle analytic? In R. Heck (ed) *Logic, Language and Thought*, Oxford University Press, Oxford. Reprinted in Boolos (1998). Page references to reprint.

Boolos, G.: 1998, *Logic, logic, and logic*, Harvard Univesity Press, Cambridge, MA.

Brandom, R.: 1998, *Making it Explicit: Reasoning, Representing, and Discursive Commitment*, Harvard University Press, Cambridge, MA.

Braüner, T. and Ghilardi, S.: 2007, First-order modal logic. In P. Blackburn, J. van Benthem and F. Wolter (eds), *Handbook of modal logic*, Elsevier, Amsterdam.

Bull, R. and Segerberg, K.: 2001, Basic modal logic. In D. Gabbay and F. Guenthner (eds) *Handbook of Philosophical Logic*, Vol. 3, 2nd edn, Kluwer, Dordrecht.

Burali-Forti, C.: 1897, Una questione sui numeri transfiniti, *Rendiconti del Circolo Matematico di Palermo* **11**(1), 154–164.

Burge, T.: 1974, Truth and singular terms, *Noûs* **8**(4), 309–325.

Burge, T.: 1979, Individualism and the mental, *Midwest studies in philosophy* **4**(1), 73–121.

Burgess, J. P.: 2002, Basic tense logic. In D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic*, Vol 7, 2nd edn, Kluwer, Dordrecht.

Burgess, J. P.: 2004, E pluribus unum: Plural logic and set theory, *Philosophia Mathematica* **12**(3), 193–221.

Burgess, J. P.: 2005, *Fixing Frege*, Princeton University Press, Princeton, NJ.

Burgess, J. P. and Rosen, G.: 1997, *A Subject with no Object*, Clarendon Press, Oxford.

Cartwright, R.: 1994, Speaking of everything, *Noûs* **28**(1), 1–20.

Church, A.: 1996, *Introduction to Mathematical Logic*, Princeton University Press, Princeton, NJ.

Corcoran, J.: 2006, Schemata: the concept of schema in the history of logic, *Bulletin of Symbolic Logic* **12**(2), 219–240.

Davidson, D.: 1967, Truth and meaning, *Synthese* **17**(1), 304–323.

Davidson, D.: 1973, Radical interpretation, *Dialectica* **27**(3-4), 313–328.

Davidson, D.: 1974a, On the very idea of a conceptual scheme, *Proceedings and Addresses of the American Philosophical Association*, Vol. 47, pp. 5–20.

Davidson, D.: 1974b, Replies to David Lewis and W. V. Quine, *Synthese* **27**(3), 345–349.

Devitt, M.: 1981, *Designation*, Columbia University Press, New York, NY.

Dummett, M.: 1963, The philosophical significance of Gödel's theorem, *Ratio* **5**(2), 140–55.

Dummett, M.: 1983, *Frege: Philosophy of Language*, 2nd edn, Duckworth, London.

Dummett, M.: 1991, *Frege: Philosophy of mathematics*, Duckworth, London.

Fine, K.: 1977, Prior on the construction of possible worlds and instants. Postscript to A. N. Prior and K. Fine, *Worlds, Times and Selves*, Duckworth, London.

Fine, K.: 1981, First-order modal theories I—sets, *Nous* **15**(2), 177–205.

Fine, K.: 2003, The problem of possibilia. In M. Loux and D. Zimmerman (eds), *The Oxford Handbook of Metaphysics*, Oxford University Press, Oxford.

Fine, K.: 2005, Our knowledge of mathematical objects. In T. S. Gendler and J Hawthorne (eds), *Oxford Studies in Epistemology*, vol. 1, Clarendon Press, Oxford.

Fine, K.: 2006, Relatively unrestricted quantification. In Rayo and Uzquiano (2006).

Fine, K.: 2007, Response to Alan Weir, *Dialectica* **61**, 117–125.

Fitting, M. and Mendelsohn, R. L.: 1998, *First-order Modal Logic*, Kluwer, Dordrecht.

Fraenkel, A., Bar-Hillel, Y. and Levy, A.: 1973, *Foundations of Set Theory*, North-Holland, Amsterdam.

Frege, G.: 1892, Uber Sinn und Bedeutung, *Zeitschrift für Philosophie und philosophische Kritik* **100**(1), 25–50. Translated as 'On Sense and Reference' by M. Black in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds and trans), Blackwell, Oxford, 1980.

Frege, G.: 1893, *Grundgesetze der Arithmetik*, Vol. I, Verlag Herman Pohle, Jena. Partially translated into English by M.Furth as *The Basic Laws of Arithmetic*, University of California Press, Berkeley, CA, 1964.

Garson, J. W.: 2001, Quantification in modal logic. In D. Gabbay and F. Guenthner (eds) *Handbook of Philosophical Logic*, Vol. 3, 2nd edn, Kluwer, Dordrecht.

Geach, P.: 1968, *Reference and Generality*, Cornell University Press, Ithaca, NY.

Glanzberg, M.: 2004, Quantification and realism, *Philosophy and Phenomenological Research* **69**(3), 541–572.

Glanzberg, M.: 2006, Context and unrestricted quantification. In Rayo and Uzquiano (2006).

Gödel, K.: 1947., What is Cantor's continuum problem, *The American Mathematical Monthly* **54**(9), 515–25.

Gödel, K.: 1972, Some remarks on the undecidability results. In his *Collected Works*, (ed) S. Feferman, Vol. 2, Oxford University Press, Oxford, 1990.

Gödel, K.: 1983, What is cantor's continuum problem (expanded version). An expanded version of Gödel (1947.). In Benacerraf and Putnam (1983).

Goldblatt, R.: 1992, *Logics of Time and Computation*, 2nd edn, CSLI, Stanford, CA.

Grandy, R.: 1973, Reference, meaning, and belief, *The Journal of Philosophy* **70**(14), 439–452.

Hallett, M.: 1984, *Cantorian Set Theory and Limitation of Size*, Clarendon Press, Oxford.

Heck, R.: 1996, The consistency of predicative fragments of Frege's Grundgesetze der Arithmetik, *History and Philosophy of Logic* **17**(1-2), 209–220.

Hellman, G.: 2006, Against 'absolutely everything'! In Rayo and Uzquiano (2006).

Hirsch, E.: 2002, Quantifier variance and realism, *Philosophical Issues* **12**(1), 51–73.

Hirsch, E.: 2005, Physical-object ontology, verbal disputes, and common sense, *Philosophy and Phenomenological Research* **70**(1), 67–97.

Hirsch, E.: 2009, Ontology and alternative languages. In D. Chalmers and D. Manley and R. Wasserman (eds), *Metametaphysics*, Oxford University Press, Oxford.

Hughes, G. E. and Cresswell, M. J.: 1996, *A New Introduction to Modal Mogic*, Routledge, New York, NY.

Jech, T.: 2003, *Set Theory*, 3rd, revised edn, Springer, New York, NY.

Kanamori, A.: 2009, *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*, Springer, New York, NY.

Kaplan, D.: 1989, Demonstratives. In J. Almog, J. Perry and H. Wettstein (eds), *Themes from Kaplan*, Oxford University Press.

Keenan, E.: 1984, A Boolean approach to semantics, *Groenendijk, Janssen, Stockhof (eds), Truth, Interpretation and Information, Foris, Dordrecht* pp. 65–98.

Keenan, E. and Stavi, J.: 1986, A semantic characterization of natural language determiners, *Linguistics and Philosophy* **9**(3), 253–326.

Kleene, S. C.: 2002, *Mathematical Logic*, Dover Publications, Mineola, NY.

Kolaitis, P. G. and Väänänen, J. A.: 1995, Generalized quantifiers and pebble games on finite structures, *Annals of Pure and Applied Logic* **74**(1), 23–75.

Kripke, S.: 1981, *Naming and Necessity*, Blackwell, Oxford.

Lambert, K.: 2001, Free logics. In L. Goble (ed) *The Blackwell guide to philosophical logic*, Blackwell, Oxford.

Lavine, S.: 2006, Something about everything: Universal quantification in the universal sense of universal quantification. In Rayo and Uzquiano (2006).

Levy, A.: 1960, Axiom schemata of strong infinity in axiomatic set theory, *Pacific Journal of Mathematics* **10**(1), 223–238.

Lewis, D.: 1974, Radical interpretation, *Synthese* **27**(3), 331–344.

Lewis, D.: 1983, New work for a theory of universals, *Australasian Journal of Philosophy* **61**(4), 343–377.

Lewis, D. K.: 1970, General semantics, *Synthese* **22**(1), 18–67.

Lewis, D. K.: 1980, Index, context, and content. In S. Kanger and S. Öhman (eds), *Philosophy and Grammar*, Reidel, Dordrecht. Reprinted in Lewis (1998). Page references to reprint.

Lewis, D. K.: 1991, *Parts of Classes*, Basil Blackwell, Oxford.

Lewis, D. K.: 1998, *Papers in Philosophical Logic*, Cambridge University Press, Cambridge.

Lindström, P.: 1966, First order predicate logic with generalized quantifiers, *Theoria* **32**, 186–195.

Linnebo, Ø.: 2006, Sets, properties, and unrestricted quantification. In Rayo and Uzquiano (2006).

Linnebo, Ø.: 2010, Pluralities and sets, *The Journal of Philosophy* **107**(3), 144–164.

Linnebo, Ø.: 2013, Plural quantification, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2013 edn.

Linnebo, Ø.: n.d., The potential hierarchy of sets. Unpublished manuscript.

Linnebo, Ø. and Nicolas, D.: 2008, Superplurals in English, *Analysis* **68**(3), 186–197.

Linnebo, Ø. and Rayo, A.: 2012, Hierarchies ontological and ideological, *Mind* . Forthcoming.

Maddy, P.: 1980, Perception and mathematical intuition, *The Philosophical Review* **89**(2), 163–196.

Mathias, A.: 2001, Slim models of Zermelo set theory, *Journal of Symbolic Logic* **66**(2), 487–496.

McGee, V.: 1997, How we learn mathematical language, *The Philosophical Review* **106**(1), 35–68.

McGee, V.: 2000, 'Everything'. In G. Sher and R. Tieszen (eds), *Between Logic and Intuition: Essays in Honor of Charles Parsons*, Cambridge University Press, Cambridge.

McGee, V.: 2006, There's a rule for everything. In Rayo and Uzquiano (2006).

Montague, R.: 1970, Universal grammar, *Theoria* **36**(3), 373–398.

Mostowski, A.: 1957, On a generalization of quantifiers, *Fundamenta Mathematicae* **44**(1), 12–36.

Negri, S.: 2005, Proof analysis in modal logic, *Journal of Philosophical Logic* **34**, 507–44.

Parsons, C.: 1974, Sets and classes, *Noûs* **8**(1), 1–12.

Parsons, C.: 1977, What is the iterative conception of set? In R. Butts and J. Hintikka (eds), *Proceedings of the 5th International Congress of Logic, Methodology and Philosophy of Science 1975, Part I: Logic, Foundations of Mathematics, and Computability Theory*, Reidel, Dordrecht. Reprinted in Benacerraf and Putnam (1983). Pages references to reprint.

Parsons, C.: 1983, Sets and modality. In his *Mathematics in Philosophy*, Cornell University Press, Ithaca, NY.

Parsons, C.: 2006, The problem of absolute universality. In Rayo and Uzquiano (2006).

Paseau, A.: 2007, Boolos on the justification of set theory, *Philosophia Mathematica* **15**(1), 30–53.

Peters, S. and Westerståhl, D.: 2006, *Quantifiers in Language and Logic*, Clarendon Press, Oxford.

Priest, G.: 2006, *In Contradiction: A Study of the Transconsistent*, Clarendon Press, Oxford.

Prior, A. N.: 1960, The runabout inference ticket, *Analysis* **21**(2), 38–39.

Putnam, H.: 1981, *Reason, Truth, and History*, Cambridge University Press, Cambridge.

Putnam, H.: 1987a, *The Many Faces of Realism*, Open Court, La Salle, IL.

Putnam, H.: 1987b, Truth and convention: On Davidson's refutation of conceptual relativism, *Dialectica* **41**(1-2), 69–77.

Putnam, H.: 2000a, Paradox revisited I: Truth. In Sher and Tieszen (2000).

Putnam, H.: 2000b, Paradox revisited II: Sets—a case of all or none. In Sher and Tieszen (2000).

Quine, W. V. O.: 1960, *Word and Object*, MIT Press, Cambridge, MA.

Quine, W. V. O.: 1973, *The Roots of Reference*, Open Court, La Salle, IL.

Quine, W. V. O.: 1974, *Methods of logic*, Routledge and Keegan Paul, London.

Quine, W. V. O.: 1981, *Mathematical Logic*, Harvard University Press, Cambridge, MA.

Rayo, A.: 2002, Word and objects, *Noûs* **36**(3), 436–464.

Rayo, A.: 2006, Beyond plurals. In Rayo and Uzquiano (2006).

Rayo, A. and Uzquiano, G.: 1999, Toward a theory of second-order consequence, *Notre Dame Journal of Formal Logic* **40**(3), 315–325.

Rayo, A. and Uzquiano, G. (eds): 2006, *Absolute Generality*, Oxford University Press, Oxford.

Rescher, N.: 1962, Plurality quantification, *Journal of Symbolic Logic* **27**(3), 373–374.

Restall, G.: 2005, Multiple conclusions. In P. Hajek, L. Valdes-Villanueva and D. Westerstahl (eds) *12th International Congress on Logic, Methodology and Philosophy of Science*, Kings' College Publications, London.

Russell, B.: 1902, Letter to Frege. In J. van Heijenoort (ed), *From Frege to Gödel*, Harvard University Press, Cambridge, MA., 1967, 124-125.

Russell, B.: 1906, The theory of implication, *American Journal of Mathematics* **28**(2), 159–202.

Shapiro, S.: 1991, *Foundations without Foundationalism: A Case for Second-order Logic*, Oxford University Press, Oxford.

Shapiro, S.: 2003, All sets great and small: And I do mean ALL, *Philosophical Perspectives* **17**(1), 467–490.

Shapiro, S. and Wright, C.: 2006, All things indefinitely extensible. In Rayo and Uzquiano (2006).

Sher, G. and Tieszen, R. (eds): 2000, *Between Logic and Intuition: Essays in Honor of Charles Parsons*, Cambridge University Press.

Shoenfield, J.: 1967, *Mathematical logic*, Addison-Wesley, Reading, MA.

Sider, T.: 2009, Ontological realism. In D. Chalmers and D. Manley and R. Wasserman (eds), *Metametaphysics*, Oxford University Press, Oxford.

Sider, T.: 2010, Against parthood. Manuscript.

Stalnaker, R.: 1977, Complex predicates, *The Monist* **60**(3), 327–339.

Stanley, J. and Szabó, Z. G.: 2000, On quantifier domain restriction, *Mind & Language* **15**(2-3), 219–261.

Stanley, J. and Williamson, T.: 1995, Quantifiers and context-dependence, *Analysis* **55**(4), 291–295.

Tarski, A.: 1933, The concept of truth in the languages of the deductive sciences, *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych* **34**, 13–172. English translation in Tarski (1983).

Tarski, A.: 1983, Logic, semantics, metamathematics. J. Corcoran (ed), Hackett Publishing Company, Indianapolis, IN.

Uzquiano, G.: 1999, Models of second-order Zermelo set theory, *Bulletin of Symbolic Logic* **5**(3), 289–302.

Uzquiano, G.: 2006, Unrestricted unrestricted quantification: The cardinal problem of absolute generality. In Rayo and Uzquiano (2006).

Uzquiano, G.: 2009, Quantification without a domain. In Ø Linnebo and O Bueno (eds) *New Waves in the Philosophy of Mathematics*, Palgrave, Basingstoke.

Venema, Y.: 2001, Temporal logic. In L. Goble (ed) *The Blackwell guide to philosophical logic*, Blackwell, Oxford.

Viganó, L.: 2000, *Labelled Non-Classical Logics*, Kluwer, Dordrecht.

von Neumann, J.: 1927, Zur Hilbertschen Beweistheorie, *Mathematische Zeitschrift* **26**(1), 46.

Wang, H.: 1974, *From Mathematics to Philosophy*, Routledge, London.

Westerståhl, D.: 1985a, Determiners and context sets. In J. van Bentham and A. ter Meulen (eds), *Generalised Quantifiers in Natural Language*, Foris, Dordrecht.

Westerståhl, D.: 1985b, Logical constants in quantifier languages, *Linguistics and Philosophy* **8**(4), 387–413.

Williamson, T.: 1998a, Bare possibilia, *Erkenntnis* **48**(2-3), 257–273.

Williamson, T.: 1998b, Indefinite extensibility. In J. Brandl and P. Sullivan (eds) *New essays on the philosophy of Michael Dummett*, Rodopi, Amsterdam.

Williamson, T.: 2003, Everything, *Philosophical Perspectives* **17**(1), 415–465.

Williamson, T.: 2004, Philosphical 'intuitions' and scepticism about judgement, *Dialectica* **58**(1), 109–153.

Williamson, T.: 2006, Absolute identity and absolute generality. In Rayo and Uzquiano (2006).

Williamson, T.: 2007, *The Philosophy of Philosophy*, Blackwell, Oxford.

Zermelo, E.: 1908, Untersuchungen über die Grundlagen der Mengenlehre. i, *Mathematische Annalen* **65**(2), 261–281. English translation in Zermelo (2009).

Zermelo, E.: 1930, Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre, *Fundamenta mathematicae* **16**, 29–47. English translation in Zermelo (2009).

Zermelo, E.: 2009, *Collected Works/Gesammelte Werke: Volume I/Band I - Set Theory, Miscellanea/Mengenlehre, Varia.* (eds) A Kanamori, A et al., Springer, New York, NY.