Part II

Chapter 6

The Language of Logical Possibility

6.1 Agenda

In Part I we discussed some existing potentialist and actualist approaches to the foundations of set theory, and noted some problems for them. We saw that actualists face an arbitrariness problem related to the Buralli-Forti paradox. Potentialists face a problem about clarifying the intended meaning their possibility and extendability claims (e.g., the intended interpretation of the $\Box$ and $\Diamond$) in a way that supports their project. And both actualists and potentialists face a problem about justifying the ZFC axioms, especially the Axiom of Replacement. With this picture in place, we can now begin this book's positive project.

In the next few chapters (Part II of this book) I will develop a particular version of (Putnamian) Potentialist set theory, and argue that it lets us avoid many of the problems discussed above. As we have seen, potentialist paraphrases of set theory make claims about how it would be (in some sense) possible to extend an initial segment of the hierarchy of sets.

So I will begin by introducing the key notion of possibility which I think we should use to formulate potentialist set theory - namely, **conditional logical possibility** - in this chapter. I will argue that we already have and need a notion of logical possibility, and that this notion naturally generalizes to a notion of what is logical possible while 'holding certain structural facts fixed'. The resulting notion of conditional logical possibility turns out to have significant expressive power. For example, combining it with first order vocabulary lets us categorically describe the structure of the natural numbers.

In chapter 7 I will discuss how we can use the relevant notion of logical possibility to formulate a version of Putnam's potentialist set theory which differs from, and simplifies, Hellman's formulation in a few key ways.

In chapter 8 I will give mathematical details on my proposed formulation. Readers who are more interested in this book's philosophical arguments than the justification for Replacement provided in the Part III, are invited to skim this section or only consult it as needed to answer questions about how the ideas in chapter 7 can be implemented.

Finally in chapter 9 I will review how this helps solve the problems raised for other versions of potentialism above.]

6.2 Logical Possibility Simpliciter

We seem to have an intuitive notion of logical possibility which applies to claims like $(\exists x)(red(x) \wedge round(x))$ and makes sentences like the following come out true.

- It is logically possible that $(\exists x)(red(x) \wedge round(x))$
- It is not logically possible that $(\exists x)(red(x) \wedge \neg red(x))$
- It is logically necessary that $(\forall x)(red(x)) \rightarrow \neg(\exists x)(\neg red(x))$.

Philosophers representing a range of different views of mathematics have made use of this notion and are comfortable applying it to non-first order sentences.

This notion of logical possibility is (roughly) what's analyzed by saying some theory has a set theoretic model. And it concerns whether some state of affairs is allowed by the most 'general subject matter neutral laws' of how there can be some pattern of objects standing in n-place relations (in something like Frege's sense of logical laws being subject matter neutral).

However, like Hartry Field, I take it logical possibility to be a primitive concept not reducible to any facts about set theoretic models or possible worlds. If you are skeptical that there is such a notion, note that it is definable in terms of the even more common notion of validity[1] (something

[1] See the discussion of the corresponding notion of validity and logical consequence in

is logically possible iff its negation is not logically necessary iff the inference from the empty premises to its negation is not valid).

At first glance, one might be tempted to identify claims about logical possibility with claims about the existence of a set theoretic model. But see [33] for a convincing argument against this view. Very crudely, the issue is this: a key aspect of our notion of logical possibility is that what's actual must be logically possible. But, if we identify logical possibility with the existence of a set theoretic model then it seems puzzling why the inference from actual to possible is justified, since the total universe can't be represented as a set theoretic model (because, assuming realism about set theory, the universe contains all the sets)[2].

To evaluate whether a claim ϕ is logically possible (in this sense), we hold fixed the operation of logical vocabulary (like $\exists, \wedge, \vee, \neg$), but abstract away from any further metaphysically necessary constraints on the application of particular relation symbols. Thus, we consider all possible ways for relations to apply (including those ways that aren't definable). For example, it is logically possible that $(\exists x)(\text{Raven}(x) \wedge \text{Vegetable}(x))$, even if it would be metaphysically impossible for anything to be both a raven and a vegetable.

We also abstract away from constraints on the size of the universe, so that

[31],[87] and [44].

[2]Of course, in the case of first order statements, Gödel's completeness theorem ensures us that any intuitively logically possible state of affairs will have a set theoretic model. But, it seems that we can understand claims about the logical possibility of states of affairs described with richer logical vocabulary like second order quantification possibility (or at least that we shouldn't rule this possibility out in advance). Also, as Field argues, our willingness to draw this conclusion from Godel's proofs in the completeness theorem seems to arise from our having an intuitive notion of logical possibility.

$\Diamond(\exists x)(\exists y)(\neg x = y)$ would be true even if the actual universe contained only a single object.

6.3 Contrast With Other Modal Notions

Before going on, it may help philosophical readers to note how my notion of logical possibility differs from three vaguely similar modal notions in the literature, namely Tarskian re-interpretability, metaphysical possibility and conceptual possibility.

The notion of logical possibility is (potentially) less demanding than the notion of Tarskian reinterpretability, for reasons discussed in Etchemendy's *The Concept Of Logical Consequence*. Essentially, the issue is that certain scenarios might be genuinely logically possible but require the existence of more objects than actually exist, and hence not permit any Tarskian reinterpretation (since Tarskian reinterpretations of a sentence must still take the sentence's quantifiers to range over some collection of objects in the actual world).

The notion of logical possibility is also prima facie less demanding than the notion of metaphysical possibility[3]. For, as Frege noted, the laws of logic hold at all possible worlds. Yet it would seem that statements like $(\exists x)(Raven(x) \wedge Vegetable(x))$ can require something which is logically pos-

[3]I want to leave it open for hardcore Tractarians who reject this notion to say that logical possibility turns out to be the same thing as metaphysical possibility. I'm just noting that we have a concept of logical possibility independent of this, and that this suffices to give an attractive account of set theory.

sible but metaphysically impossible.

Cashing set theory out in terms of logical rather than metaphysical possibility frees us from worries that there might be overly restrictive limits on the number of physical objects (and perhaps the number of impure mathematical objects) arising from constraints on the structure of space and time[75].

Finally, the notion of logical possibility is also less demanding than the notion(s) of idealized conceivability and/or conceptual possibility at issue in debates over philosophical zombies and Chalmers' *Constructing the World*[16] (and are, inconveniently, sometimes also labeled logical possibility). For the notion of conceptual possibility reflects something like ideal a priori acceptability. So when evaluating whether it is conceptually possible that ϕ we have to preserve all analytic truths associated with relations occurring in ϕ. In contrast (as I have noted above) logical possibility abstracts away from all such specific features of relations. Thus, for example, if we assume it is analytic that $(\forall x)(bachelor(x) \rightarrow male(x))$, then it will be logically possible but *not* conceptually possible that $(\exists x)(bachelor(x) \wedge \neg male(x))$.

6.4 Conditional Logical Possibility

So much for the familiar notion of logical possibility discussed above, and the appeals of taking it as a conceptual primitive. I want to suggest that, if you accept this notion, it's natural to also accept a generalization of it which I will call **conditional** logical possibility.

If you accept a notion of logical possibility, it seems only natural to be able to make sense of restricting that notion to the scenarios which preserve the structure of how some relations $R_1, \ldots, R_n$ apply (in the actual world, or in some possible world under consideration).

So I think we can also intuitively understand claims about logical possibility given the facts about how certain relations apply. Consider a statement like the following.

> Given what cats and blankets there are, it is logically impossible that each cat slept on a different blanket last night.

This sentence has an intuitive reading which employs a notion of logical possibility *holding fixed the way that certain relations apply* (in this case, holding fixed what cats and blankets there are) rather than logical possibility *simpliciter*. A moment's thought will reveal that (on this reading) the above sentence is true if and only if there are more cats than blankets.

I propose to think of the logical possibility $\Diamond_{(\ldots)}(\ldots)$ as an operator which takes a sentence ϕ and a finite (potentially empty) list of relation symbols $R_1, \ldots, R_n$ and produces a sentence $\Diamond_{R_1,\ldots,R_n}\phi$ which says that it is logically possible for ϕ to be true, without any change to (structural facts) about how the relations $R_1, \ldots R_n$ apply. Thus, for example, the claim, 'Given what cats and baskets there are, it is logically impossible that each cat slept in a distinct basket' becomes:

$$\mathbf{C}\wedge\mathbf{B}{:}\neg\Diamond_{cat,basket}[(\forall x)(cat(x) \to (\exists y)(basket(y) \wedge sleptIn(x,y) \wedge (\forall z)[cat(z) \wedge sleptIn(z,y) \to x = z])]$$

Stepping back, we might say that claims employing the familiar the notion of logical possibility simpliciter ($\Diamond$) discussed in the previous sections concern what's possible if we let both the size of the domain of discourse and the application of relations to that domain vary with complete freedom. In contrast, claims about conditional logical possibility ($\Diamond_{R_1,...R_n}$) concern what's logical possible if we hold fixed the structural facts about how some relations $R_1, \ldots, R_n$ apply (while still letting the size of the domain and the application of all relations other than $R_1, \ldots, R_n$ within this domain vary freely).

But what do I mean by 'holding the structural facts fixed', about how some relations $R_1, \ldots, R_n$ apply? Consider the way we'd say two different interpretations of some person's language which both take the 'number' and 'successor' to apply to some ω sequence thereby agree on the *structure* of the natural numbers (under successor), even if they disagree about the total size of the universe, whether Julius Caesar or the empty set are identical to any numbers etc. Or consider the way that a Platonist would say the structure of the natural numbers is fixed necessarily and will always remain the same, even if the total size of the universe can be changed by the creation or destruction of physical objects or changes to the structure of space etc.

My notion of agreeing on/preserving structural facts about $R_1, \ldots, R_n$ generalizes this notion of agreeing on/preserving the structural facts about how 'natural number' and 'successor' apply.

If we could talk about functions between (the objects in) different logically possible worlds, then we could specify what it takes to hold the structural facts about how some relation (say, admires()) apply fixed, in terms of isomorphisms as follows.

> A logically possible world w_2 counts as holding fixed the structural facts about how admires() applies in w_1 iff some function f bijectively maps the objects which either admire or are admired in w_1 to the objects which either admire or are admired in w_2, so that for all objects x and y in w_1 which either admire or are admired in w_1, we have x admires y iff $f(x)$ admires $f(y)$.
>
> And more generally, a logically possible world w_2 preserves the structural facts about how relations $R_1, \ldots, R_n$ (say admires() and cat()) apply iff some function f bijectively maps the objects which $R_1, \ldots, R_n$ apply to in w_1 (i.e. those things which are either cats or admire something or are admired by something) to the objects which $R_1, \ldots, R_n$ apply to in w_2 in a way that respects all these relations.

But note that I'm putting conditional logical possibility forward as a conceptual and metaphysical primitive. So I make this comparison purely for expository purposes.

For another example of how conditional logical possibility applies, consider the claim that it's logically possible, holding fixed how 'admires()' applies, that everything which admires something it is pitied by something. That is,

$$\Diamond_{\text{admires}}(\forall x)\left[(\exists y)\text{admires}(x,y) \rightarrow (\exists z)\text{pities}(z,x)\right]$$

.

In contrast to the bare logical possibility claim,

$$\Diamond(\forall x)(\exists y)\,(\text{admires}(x, y) \rightarrow (\exists z)\text{pities}(z, x))$$

we can't see that this conditional logical possibility $\Diamond_{\text{admires}}$ claim is true just by considering a logically possible scenario in which a single individual both admires and pities themselves. For presumably the actual world contains more than one individual who admires and/or is admired by something. Thus, this scenario will not qualify as holding the structural facts about how admiration() applies fixed since it has less individuals who admire or are admired than the actual world.

Instead, to witness the truth of this claim, we must consider a logically possible scenario which is 'isomorphic to reality' with regard to which objects admire one another, in the sense described above. So, for example, such scenarios cannot change the number of individuals which admire something, or the number of individuals who admire themselves, or the number of pairs of individuals who reciprocally admire each other. But this logically possible scenario can vary the application of other relations like pities(), (and, though not needed here, the total size of the universe) as needed to make the claim $(\forall x)\,((\exists y)\text{admires}(x, y) \rightarrow (\exists z)\text{pities}(z, x))$ true. Thus, for example, to witness the truth of the above logical possibility claim, we can consider a scenarios where for all a and b, if a admires b, b pities a (even though one hopes is not true of the actual world).

Note that fixing the structural facts about how some relations $R_1, \ldots R_n$

apply doesn't merely mean preserving the truthvalue of some set of sentences describing how these relations apply. For example, suppose the ticks of some infinite clock form a genuine ω sequence under the relation 'follows'. Then when evaluating $\Diamond_{tick,follows}$ only logically possible scenarios which preserve the fact that the clock ticks form an ω sequence under 'follows' will be relevant. Yet by Skolemization no collection of first order sentences will suffice to ensure that this is the case. Thus, the claim that some logically possible scenario preserves the structural facts about how 'tick' and 'follows' apply in the actual world (or some scenario under consideration) cannot be identified with the requirement that this scenario preserves the truthvalue of some collection of first order sentences.

4

We can also make *nested* logical possibility claims, i.e., claims about the

[4]One can further explain and motivate my the notion of conditional logical possibility by relating it to Stuart Shapiro's notion of structures in [98]. Shapiro introduces a notion of systems, consisting of some objects to which some relations $R_1...R_n$ apply, considered under some relations. He gives the following examples, "An extended family is a system of people with blood and marital relationships, a chess configuration is a system of pieces under spatial and "possible move relationships, a symphony is a system of tones under temporal and harmonic relationships, and a baseball defense is a collection of people with on-field spatial and "defensive-role relations."

Then he says that a structures are 'the abstract form' of a system, which we get by "highlighting the interrelationships among the objects, and ignoring any features of them that do not affect how they relate to other objects in the system." Thus, for example, "The natural-number structure is exemplified by the strings on a finite alphabet in lexical order, an infinite sequence of strokes, an infinite sequence of distinct moments of time, and so on." And adding or subtracting objects to the world outside of a given system, will make no different to which structure that system instantiates.

Although I mean to propose them as conceptual primitives, one can (roughly) explain my notion of conditional logical possibility (aka structure preserving logical possibility) in terms of Shapiro's notions as follows:

It is logical possible, given the $R_1 \ldots R_n$ facts, that ϕ (i.e., $\Diamond_{R_1 \ldots R_n}$ iff some logically possible scenario makes ϕ true while holding fixed what structure (in Shapiro's sense) the system formed by the objects related by $R_1 \ldots R_n$ (considered under the relations $R_1 \ldots R_n$) instantiates.

logical possibility of scenarios which are themselves described in terms of logical possibility. I have in mind sentences like the following:

> $\Diamond$**C**$\wedge$**B**: $\Diamond(\neg\Diamond_{cat,basket}\,[(\forall x)(cat(x) \rightarrow (\exists y)(basket(y) \wedge sleptIn(x,y) \wedge (\forall z)[cat(z) \wedge sleptIn(z,y) \rightarrow x = z])])$

The above sentence, $\Diamond$(C$\wedge$B), expresses a truth because (reading from the outside in):

- It is logically possible (holding fixed nothing) that there are 4 cats and 3 baskets.
- Relative to the logically possible scenario where there are 4 cats and 3 baskets, it is not logically possible (given what cats and baskets there are), that each cat slept in a basket and no two cats slept in the same basket.

Based on these kind of examples, I take logical possibility sentences of the form $\Diamond_{R_1,\ldots,R_n}\phi$ to be meaningful, even in cases where ϕ is itself a sentence which makes appeal to facts about logical possibility. As noted above, I will not allow sentences which quantify in to the $\Diamond$ of logical possibility, i.e. sentences of the form $(\exists x)\Diamond\phi(x)$.

To clearly express claims about logical possibility, we can define a formal language, which I will call the language of logical possibility (though no implication that this exhausts the concept of logical possibility should be drawn). Fix some infinite collection of variables and relation symbols of every arity together with $\perp$ and define the language of logical possibility to be the smallest language built from these variables using these relation symbols and equality closed under applications of the normal first order connectives and quantifiers and $\Diamond_{...}$ (where $\Diamond_{...}$ expressions can only be applied to sentences (so there is no quantifying in). We will also use $\Box_{...}$ in our sentences but regard it as an abbreviation for $\neg\Diamond_{...}\neg$.

In appendix B I show how to use set theoretic models to approximate truth conditions for arbitrary sentences in the language of logical possibility. Looking at this appendix may help answer any questions about the meaning of nested conditional logical possibility claims.

6.5 A Sample of this Language's Power

To see the expressive power of this notion of logical possibility (and begin to see how it can be useful for formulating potentialist set theory) note that we can categorically describe the intended structure of the natural numbers, using only the conditional logical possibility operator $\Diamond_{...}$ and standard first order logical connectives, rather than second order quantification (as is commonly done).

One can uniquely describe the intended structure of the natural numbers by combining the first 6 Peano Axioms (which can be expressed using only first order logical vocabulary) with an Axiom of Induction, which can be expressed in the language of second order logic as follows:

$$(\forall X)(([0 \in X \wedge (\forall n)(n \in X \rightarrow S(n) \in X)] \rightarrow (\forall n)(\mathbb{N}(n) \rightarrow n \in X))$$ [5]

[5] Where 0 is not officially part of our language, but I use claims about 0 to abbreviate corresponding claims about the unique number that isn't a successor of anything, in the

Informally, this axiom says that if some property applies to 0 and to the successor of every number it applies to, then it applies to all the numbers.

We can express the same idea using $\Diamond$ (and any one place relation P other than '$\mathbb{N}$') as follows.[6]

$$\Box_{\mathbb{N},S}\Big([P(0) \wedge (\forall x)(\forall y)(P(x) \wedge S(x,y) \rightarrow P(y))] \rightarrow (\forall x)(\mathbb{N}(x) \rightarrow P(x))\Big)$$

This formula says that, given the facts about what is a number and a successor, i.e., about how $\mathbb{N}$ and S apply), it would be logically impossible for P to apply to 0 and to the successor of each object which it applies to without applying to all the numbers.

Call the sentence you get by replacing the axiom of induction in second order Peano Arithmetic with the above modal sentence $PA_{\Diamond}$.

usual fashion.

[6] Here $P(0)$ is shorthand for $(\exists z)(\forall w)\,(\mathbb{N}(z) \wedge \neg S(w,z) \wedge P(z))$, and $\Box$ abbreviates $\neg\Diamond\neg$ as usual.

Chapter 7

Purified Potentialist Set Theory: An Informal Sketch

In this chapter I will informally present my preferred version of potentialist set theory (using the notion of conditional logical possibility), and clarify some philosophical issues about it. This formulation of potentialist set theory will avoid appeal to the contested notion of *de re* modality and (arguably) uses fewer conceptual primitives than other articulations of potentialism. In later chapters I will argue that it solves some of the problems noted for these previous views.

Structurally my version of potentialst paraphrases (presented in section 8) will look rather like Hellman's paraphrases in [44]. Recall that Hellman would paraphrase the set theoretic sentence $(\forall x)(\exists y)(x \in y)$ as saying the following. Necessarily if V_1 satisfies ZFC_2 and includes a set x, it is logically possible for there to be an extension V_2 of V_1, also satisfying ZFC_2 and containing a set y such that $x \in y$ (in the sense of $\in$ relevant to V_2). Writing this out formally using $\geq$ to say that one model of ZFC_2 extends another, we get

$$\Box(\forall V_1)(\forall x)\,[x \in V_1 \rightarrow \Diamond(\exists V_2)(\exists y)(y \in V_2 \wedge V_2 \geq V_1, \wedge x \in y)]$$

My paraphrases will have broadly the same structure, however there are a number of important differences to be considered.

7.1 Which Hierarchies?

First, we can ask what structure the standard width initial segments V discussed in our potentialist paraphrase are required to have. As we saw, Hellman considers the possibility of structures satisfying a version of ZFC_2.

In contrast, I will formulate potentialist set theory in terms of the possibility of structures (called initial segments) which merely satisfy the width requirements for the iterative hierarchy of sets. Thus, from an actualist point of view, any V_α (i.e., the sets below any ordinal α) will count as an initial segment. I won't even require the structures I use in my paraphrases to satisfy Boolos' very modest height requirement, that there not be a highest level.

I think this way of proceeding has a kind of intrinsic elegance, and Linnebo (in effect) takes a similar approach to this question. Moreover, by not requiring the structures in my paraphrases to satisfy ZFC_2 I need not assume the consistency of an infinite number of inaccessible as Hellman must. As we will see this allows me to justify my potentialism from a single coherent conception of logical possibility without the kind of controversial assumption which would be needed to justify the consistency of this kind of large cardinal axiom.

7.2 A Putnamian Approach to Initial Segments

Now let's consider the sense in which my potentialist paraphrases will say that objects with the structure above are possible. We saw that Putnam, Hellman and Linnebo all took very different approaches to this question. Putnam considered the possibility of first order states of affairs involving non-mathematical relations. Hellman considered the possibility of second order items X and f picking out a structure. And Linnebo considered the interpretational possibility of there being sets up to a certain height.

I will employ a version of Puntam's approach. When I say that it would be possible to have an initial segment V, I will mean (something like) that it's logically possible for some non-mathematical relations (e.g., 'is a penciled point', 'is connected by an arrow') to apply to objects with the intended structure of an iterative hierarchy of sets, in the same way that 'set' and '...is an element of...' are supposed to apply within a standard with initial segment (in the sense just introduced)[1]. By using non-mathematical rela-

[1] So, for example, although I may casually talk about the possible existence of iterative hierarchy structures V_i, I don't mean to assert that there are (or could be) special objects called structures, as e.g. Shapiro does. Or at least, I don't want to say that we need such objects to understand set theory. All talk about 'the possibility of a structure existing',

tions we avoid having to presume there is an antecedently meaningful notion of set or other mathematical relation.

But how exactly are we to cash out the claim that some relations apply to objects in a fashion which makes them an initial segment in the sense above? One might take the arrows and pencil points to form an initial segment just if they satisfy some set-theoretic axiomitization an initial segment with 'set' replaced with pencil point and 'element of' replaced with arrows. This is essentially what Hellman does (replacing the 'set' and 'in' in ZFC_2 with second order items X and f).

However, it will be convenient to admit the levels in this hierarchy as primitive objects in their own right rather than rely on the (non-obvious) fact that Von Neumann ordinals can serve that function inside the sets. So I will follow Boolos' approach in (discussed in section 2.1) by taking iterative hierarchies to involve of two different types of (first order) objects : sets and ordinal levels, with sets being related to one another by elementhood, ordinal levels being related to one another by less than, and every set being 'available at' some ordinal level.

Thus, I will employ five non-mathematical relations to characterize the notion of initial segment: two one place relations playing the role of set() and ord(), and three two place relations playing the roles of $\in$, $<$ (ordinal ordering) and @ (is available at). Note that here I am taking ord() to apply to ordinal levels, i.e., to the layers of the hierarchy of sets at which various

in the potentialist paraphrase strategy above is merely shorthand for claims about the possibility of there being objects which instantiate specific non-mathematical first order predicates and relations in a certain way.

sets are 'available', not to the sets which we'd call an ordinal in standard set theory (I will still refer to these objects as ordinals). So, for example, we might use the following non-mathematical first order properties and relations: ... is a penciled point, ...is a penciled star, ...is connected to.. by a dotted/dashed/solid arrow.

7.3 Characterizing Iterative Hierarchies Modally

I will define a formula in the language of logical possibility $\mathcal{V}(\text{set}, \text{ord}, <, \in, @)$ which asserts[2] the relations $\text{set}, \text{ord}, <, \in, @$ apply to a objects in such a way as to form a standard width initial segment as discussed above.

I will now describe the key ideas asserted by this formula, leaving the details of implementing this vision for chapter 8. The goal here is merely to convey to the reader how my notion of conditional logical possibility can capture the key ideas with minimal mathematical formalism. The reader who prefers to simply parse the mathematical definitions can jump ahead to chapter 8.

7.3.1 Describing the Levels

First we want to say that the objects playing the role of levels of the hierarchy of sets are well ordered by $<$. The claim that $<$ *linearly* orders ord is a purely first order claim, and can be expressed straightforwardly (see appendix A). However, expressing the fact that any non-empty collection of

[2]Note that although I'm choosing mathematical sounding variable names for the relations in this example, this is just a mnemonic device!

ordinals (objects satisfying ord) have a least element seem to require second order quantification. However, we can express it using conditional logical possibility.

Note that it is straightforward to express the claim that if there is a Happy ordinal there is a least Happy ordinal. We just assert that if there is a Happy ordinal there is some Happy ordinal y and that any other Happy ordinal z also is less than or equal to y. We can express this as a formula ϕ

$$\phi \overset{\text{def}}{\leftrightarrow} \exists x(\text{ord}(x) \land \text{Happy}(x)) \rightarrow \forall z\, (\text{ord}(z) \land \text{Happy}(z) \rightarrow (y \leq z))$$

.

The difficulty is to ensure that this property holds not just for the actual relation Happy but for any logically possible way a relation might apply to the ordinals. We couldn't express this property with normal logical possibility but this is exactly the content of asserting that it's logically necessary, holding fixed ord, $<$, that ϕ holds, i.e. $\Box_{<,\text{ord}}\phi$. This tells us that any logically possible way Happy applies it has the property that if there is a Happy ordinal there is a least Happy ordinal.

See A.0.2 in appendix A for details of this definition of well-ordering.

7.3.2 Describing the Sets

After requiring that the layers ord be well ordered by $<$ we then build up the sets in terms of these layers. We justify calling the objects satisfying set sets by insisting they satisfy extensionality, i.e., if $x \neq y$ are sets then they must have different elements under $\in$ and insist that the elements of a set are all sets. In accordance with the iterative hierarchy conception we insist that each set x is available at some ordinal level o, i.e. $@(x, o)$ and that a set available at o is available at ever greater ordinal[3].

This leaves the key detail of ensuring that the sets satisfy the width requirement. We do this by imposing a fatness requirement which says that, for any way we could choose some sets available at levels $o' < o$, there is a set available at level o. This is the only non-first-order aspect of the definition of $\mathcal{V}(\text{set}, \text{ord}, \in, <, @)$ aside from the definition of well-order.

As in our definition of well-order we can express the fact that (for any ordinal o) if $Happy()$ applies only to sets available at an ordinal less than o then there is a set available at o with exactly the Happy sets as elements with a first order formula[4] ψ.

As in the definition of well-order we can then use a logical necessity operator to ensure this property holds for any logically possibly way Happy could apply by asserting $\Box_{\text{set},\text{ord},<,\in,@}\psi$.

But instead we will say that it is necessary (fixing the structural facts about

[3]We also impose a thinness requirement ensuring that a set available at level o only has members available at earlier ordinals.

[4]In particular,

$$\psi \overset{\text{def}}{\leftrightarrow} (\forall o \mid \text{ord}(o)) \Big[(\forall x)(\text{Happy}(x) \to \text{set}(x) \land (\exists o')(\text{ord}(o') \land o' < o \land @(x, o'))) \to (\exists y)(@(y, o) \land (\forall z)(\text{Happy}(z) \leftrightarrow z \in y))\Big]$$

how all our relations $\text{set}, \text{ord}, <, \in, @$ apply) that, for each ord l if the property $Happy()$ only applies to sets available @ some ords below l, then there's a *set* y such that @ o which contains as elements exactly the things which are $Happy()$.

Using these techniques we can define the sentence $\mathcal{V}(\text{set}, \text{ord}, <, \in, @)$ in the language of logical possibility which says these properties apply to objects in the way that we expect the properties of 'level', 'below', 'set', 'element of' and 'available at' to apply within a standard width initial segment. Note that we can substitute any relations into this formula and remember that $\text{set}, \text{ord}, <, \in, @$ are merely placeholders for some non-mathematical properties.

We note that we can also define the notion of one initial segment extending another (in the intuitive sense where one initial segment of the sets can extend another) which denote by $V' \geq V$ where V abbreviates a list of relations $\text{set}, \text{ord}, <, \in, @$ and V' abbreviates $\text{set}', \text{ord}', <', \in'$ @.

7.4 Assignment Relations and Avoiding Quantifying In

7.4.1 Two roles for quantifying in

Now let's talk about quantifying in and how my formulation of potentialist set theory avoids it. Let us first distinguish two 'jobs' which quantifying in can do in potentialist set theory.

In general, potentialists understand set theoretic claims with nested quantification like $(\exists x)(\forall y)\neg y \in x$, as saying something like the following. It's possible to choose an x within an initial segment V, such that it is necessary that, for any y chosen from an initial segment V' extending V, $\neg y \in' x$ (where $\in'$ is the membership relation in V'). We saw that Linnebo and Hellman both use quantifying in to express this idea.

Hellman, in particular, uses two kinds of quantifying in. He uses second order quantifying in to talk how a given initial segment V could be extended (for remember that his initial segments consist of a second order object X and relation f) . And he uses first order quantifying in to make claims like 'given a set x in V it is logically possible to have an initial segment V' extending V containing a set y bearing some relation to x.'

My final task in this chapter will be to explain how both kinds of quantifying in can be eliminated.

7.4.2 Preserving the Structure of Initial Segments

Let's start with second order quantifying in. I need to do the work that Hellman does with second order quantifying-in without it, i.e., express the property that it's logically possible that an initial segment V is extended by an initial segment V'. This is done straightforwardly by simple holding fixed (the relations in) V. For instance, I could express the logical possibility of an initial segment V' extending V by

$$\Diamond_V V' \geq V \overset{\text{def}}{\leftrightarrow} \Diamond_{\text{set,ord},\in,<,@} V' \geq V$$

Now let's turn to first order quantifying in (a device which Linnebo and Hellman both use). Since we can't quantify in we must somehow pass the same information by holding fixed certain relations. In particular, to give a potentialist translation of a sentence like $(\forall x)(\exists y)\phi(x,y)$ where ϕ is quantifier free we want to assert something like the following

$$\Box\Big[\mathcal{V}(V) \rightarrow (\forall x \mid \text{set}(x)) \Diamond_V V' \geq V \wedge \big(\exists y \mid \text{set}'(y)\big)\, \phi(x,y)\Big]$$

That is, it's logically necessary however x one chooses a set x from an initial segment V that it's logically possible to extend V with an initial segment V' containing a set x making $\phi(x,y)$ true. We need to convey the same content without quantifying in. We do this by replacing quantification with an assignment function.

In particular, we associate each initial segment V with an assignment function ρ assigning variables to sets in that initial segment. Using functional notation for ρ, i.e., writing $\rho(x) = y$ rather than $\rho(x,y)$, we can capture the same content as the above sentence as follows.

$$\Box\Big[\mathcal{V}(V) \wedge \text{set}\,(\rho(\ulcorner x \urcorner)) \rightarrow \Diamond_{V,\rho,\mathbb{N},S} V' \geq V \wedge \text{set}'\,\big(\rho'(\ulcorner y \urcorner)\big)\, \phi(\rho(\ulcorner x \urcorner), \rho'(\ulcorner y \urcorner))\Big]$$

Here $\ulcorner x \urcorner$ is a number from $\mathbb{N}$ [5]) coding the variable x (see next chapter for details). Ignoring the details for the moment the key insight here is that the first logical necessity operator lets ρ range over all possible relations so the

[5] $\ulcorner x \urcorner$ is represented as $S(S(S(\ldots S(0)))$ for some number of successor operators and 0 s the unique element of $\mathbb{N}$ that isn't a successor and S is a relation that we write functionally.

consequent must be true given any choice of a set from V chosen by $\rho(\ulcorner x \urcorner)$. Similarly, the $\Diamond_{V,\rho,\mathbb{N},S}$

7.5 Big Picture

So, zooming out, (and putting these pieces together) my preferred translations of potentialist set theory have the following large scale shape.

I will translate a set theory sentence $\exists x\phi$ (where ϕ is a quantifier free formula, like $x = x$) as saying that it's logically possible that there's some structure consisting of: an initial segment of the hierarchy of sets V, a copy of the natural numbers $\mathbb{N}$ and assignment relation $\rho : \mathbb{N} \to V$ which pairs (the Gödel number of) 'x' with an object x within the hierarchy, such that $\phi(x)$.

And I will translate a set theoretic sentence $\exists x\forall y\phi(x, y)$ (where ϕ is quantifier free) roughly like this. It's logically possible that there are objects with the structure of: an initial segment V, a copy of the natural numbers $\mathbb{N}$ and assignment relation $\rho : \mathbb{N} \to V$ which pairs (the gödel number of) 'x' with an object in V so that the following claim is logically necessary holding fixed the (structural) facts about this first hierarchy $V, \rho, \mathbb{N}, S$.

> If V', ρ' consists of an initial segment V' which extends V and an assignment relation $\rho' : \mathbb{N} \to V'$ such that $\rho'(\ulcorner x \urcorner) = \rho(\ulcorner x \urcorner)$, ρ' will assign 'y' such that $\phi(\rho'(\ulcorner x \urcorner), \rho'(\ulcorner y \urcorner))$.

Chapter 8

Logico-Structural Potentialism

In this chapter I will fill in the formal details of the potentialist translation strategy described in chapter 7.

In section 8.2 I will define the core kind of structures (standard width initial segments of the hierarchy of sets) which my potentialist set theory considers the possibility of extending. In sections 8.3 and 8.4 and I will show how to cash out claims about it being possible to extend one initial segment V (and choice of some objects x, y, z within that V) to a larger V' containing an object w, without quantifying in. In section 8.5 I show how to put all the above pieces together, and provide a recursive definition for my potentialist paraphrases.

8.1 Functional Notation

To avoid overwhelming complexity we will occasionally resort to functional notation. We now explain how to understand this notation in terms of the language of logical possibility described above.

Definition 8.1.1. R is a function if $(\forall x)(\forall y)(\forall z)(R(x,y) \wedge R(x,z) \rightarrow y = z)$

So, for example, I would say that admiration is a function (in the actual world) if and only if no one admires two different people. For further notational convenience we will write $f(x_0, \ldots, x_n) = y$ to abbreviate the claim $f(x_0, \ldots, x_n, y)$. When written informally we will understand $f(x_0, \ldots, x_n)$ to stand in for some y such that $f(x_0, \ldots, x_n, y)$.

Finally, given two predicates D and R, I will say f is a function from D to R just if f is a function and $(\forall x)(D(x) \rightarrow (\exists y) f(x,y)) \wedge (\forall x \mid D(x))(\forall y)(f(x,y) \rightarrow R(y))$ and take notions surjective, injective, domain and range have their usual meaning.

8.2 Describing Standard-Width Initial Segments

So let us begin by describing our intended initial segments. Recall the iterative hierarchy conception of sets from chapter 2. Following Boolos[11] we imagined a hierarchy of sets consisting of a two sorted structure consisting of

- a well ordered series of stages, with no last element, and

- a collection of sets formed at these stages, such that a set is formed at a stage iff its members are all formed at earlier stages.

.

And we can say that something counts as a standard width initial segment if it satisfies all the requirements above except for the height requirement (that there is no last/highest stage).

I will define a formula $\mathcal{V}(\text{set}, \text{ord}, <, \in, @)$ in the language of logical possibility which expresses the fact that the objects satisfying ord are well-ordered by $<$ (giving the well-ordered series of stages), the objects satisfying set act like sets under $\in$ and that the relation @ relates sets to the stages that the are formed at in such a way that $@(x, o)$ holds just if the members of x are all available at stages before o. Note that I will sometimes write the relations $<$ and @ in infix notation, e.g., $x < y$ rather than $<(x, y)$. Also I will refer to the elements satisfying ord as ordinals and the elements satisfying set as sets.

Remember that here that I am using the terms $\text{set}, \text{ord}, \in$ for menmonic and readability purposes alone. As per the Putnamian strategy discussed in 7, my official potentialist translation of set theory will employ only logical vocabulary and non-mathematical relations like ‘is a penciled point’, and ‘there is an arrow from... to...’.

Definition 8.2.1 (Initial Segment)**.** The tuple $(\text{set}, \text{ord}, <, \in, @)$ forms an initial segment just if all of the following hold

1. $(\mathrm{ord}, <)$ is a well-order[1]

2. $(\forall x)(\forall y)\,[x \in y \to \mathrm{set}(x) \wedge \mathrm{set}(y)]$

3. $(\forall x)(\forall y)\,[@(x, y) \to \mathrm{set}(x) \wedge \mathrm{ord}(y)]$

4. $(\forall x)(\forall y)\,[x < y \to \mathrm{ord}(x) \wedge \mathrm{ord}(y)]$

5. (Fatness) For each o satisfying ord and each way of choosing some elements satisfying set from the sets (i.e. elements satisfying set) available at stages $o' < o$ there is a set with exactly those elements as members.

$$\Box_{\mathrm{set},\mathrm{ord},<,\in,@}(\forall o)\Big[\mathrm{ord}(o) \to (\forall x)\Big(P(x) \to \mathrm{set}(x) \wedge (\exists o')(\mathrm{ord}(o') \wedge o' < o \wedge @(x, o'))\Big) \to (\exists y)(\mathrm{set}(y) \wedge @(y, o) \wedge (\forall z)(P(z) \leftrightarrow z \in y)\Big)\Big]$$

6. Every set is available at some ordinal level.

$$(\forall x)[\mathrm{set}(x) \to (\exists o)\,\mathrm{ord}(o) \wedge @(x, o)]$$

7. All sets available at some o such that $\mathrm{ord}(o)$ can *only* have elements which occur at some level below as elements.

$$(\forall x)(\forall o)(@(x, o) \to (\forall z)\left[z \in x \to (\exists o')(o' < o \wedge @(z, o'))\right])$$

[1] See definition A.0.2 in appendix A for formal definition.

8. (Extensionality) No two distinct elements satisfying set have exactly the same elements.

$$(\forall x)(\forall y)\,[\mathrm{set}(x) \wedge \mathrm{set}(y) \rightarrow x = y \vee (\exists z)(\mathrm{set}(z) \wedge \neg(z \in x \leftrightarrow z \in y)]$$

9. The ordinals are disjoint from the sets

$$(\forall x)\neg(\mathrm{ord}(x) \wedge \mathrm{set}(x))$$

Note that we can think of V_α from the standard set-theoretic hierarchy as corresponding to an initial segment $\langle \mathrm{set}, \mathrm{ord}, <, \in, @\rangle$ where $\mathrm{ord}, <$ has the same order type as α. Speaking loosely, this means if $@(x, u)$ then x would be in $V_{|u|+1}$ where $|u|$ indicates the ordinal corresponding to u.

I will use $\mathcal{V}(V_i)$ to abbreviate the claim that $\mathrm{set}_i, \in_i$ etc. satisfy the sentence $\mathcal{V}(set, ord, <, \in @\rangle)$ defined above. I will also use $V(x)$ to abbreviate $\mathrm{ord}(x) \vee \mathrm{set}(x)$ and $\Diamond_V$ abbreviates $\Diamond_{\mathrm{ord},\mathrm{set},<,\in,@}$.

8.3 Extendablity

Next we need to cash out claims about one initial segment extending another.

Definition 8.3.1 (Initial Segment Extension)**.** V_a extends V_b just if all the following hold.

- $\mathcal{V}(V_a)$
- $\mathcal{V}(V_b)$

- $(\forall x)\,[\mathrm{set}_b(x) \rightarrow \mathrm{set}_a(x)]$
- $(\forall x)(\forall y)\,[\mathrm{set}_b(y) \rightarrow (x \in_b y \leftrightarrow x \in_a y)]$
- $(\forall x)[\mathrm{ord}_b(x) \rightarrow \mathrm{ord}_a(x)]$
- $(\forall x)(\forall y)\,[\mathrm{ord}_b(y) \rightarrow (x <_b y \leftrightarrow x <_a y)]$
- $(\forall x)(\forall y)[\mathrm{ord}_b(y) \rightarrow (x @_b y \leftrightarrow x @_a y)]$

I will use $V_a \geq V_b$ to abbreviate the claim that V_a (i.e. $\mathrm{set}_a, \mathrm{ord}_a, \in_a, @_a, \leq_a$) extends V_b (i.e. $\mathrm{set}_b, \mathrm{ord}_b, \in_b, @_b, <_b$).

If we followed Putnam and Hellman in quantifying in to the $\Diamond$ of logical possibility, this would suffice to let us write potentialist translations. We would translate the set theoretic utterance $(\exists x)(\forall y)\,[\neg x = y \vee \neg y \in x]$ as follows:

$$\Diamond((\exists x)\mathscr{V}(V_1) \wedge \mathrm{set}_1(x) \wedge \Box_{V_1}(\forall y)\Big[V_2 \geq V_1 \wedge \mathrm{set}_2(y) \rightarrow \neg x = y \vee \neg y \in_2 x)\Big])$$

In words it's logically possible there is an initial segment (of the hierarchy of sets) V_1 containing a set x (i.e. $\mathrm{set}_1(x)$) such that its necessary, holding fixed V_1 (i.e. $\mathrm{set}_1, \mathrm{ord}_1, \in_1, <_1, @_1$), that any choice of an element y from a model of set theory V_2 extending V_1 must satisfy $x = y \vee \neg y \in_2 x$.

However, once we embrace the notion of conditional logical possibility we can banish quantifying-in from our translations and thus avoid certain philosophical controversies (as discussed below in chapter 9).

8.4 Eliminating quantifying in

The key 'trick' which lets us eliminate quantifying in from our potentialist paraphrases, will be to supplement out initial segments V_i with a copy of the natural numbers (representing formal variables from the language of set theory) and an assignment function ρ_i which assigns each formal variable (i.e. natural number) to a set (objects satisfying set_i) from V_i. Note that my only reason for using $\mathbb{N}$ is that the natural numbers (under successor) contain infinitely many definable objects, which we can use to represent variables.

Specifically we represent the natural numbers with the predicate $\mathbb{N}$ and the function S and identify the formal variable x_n with the natural number n (i.e. $\underbrace{S(\ldots S(0)}_{n})$). Rather than use clunky formal variables x_i everywhere we instead use normal letter variables x, y, z etc.. to stand in for particular formal variables and denote the number y stands in for by $\ulcorner y \urcorner$, i.e., if y stands in for x_n then $\ulcorner y \urcorner = n$. We formalize this as follows.

Definition 8.4.1 (Interpreted Initial Segment)**.** Say that the relations in the pair (V, ρ) apply to an *interpreted initial segment* (written $\vec{\mathscr{V}}(V, \rho)$) just if $\text{set}, \in, \text{ord}, <, @$ satisfy $\mathcal{V}(\text{set}, \in, \text{ord}, <, @)$ and ρ is a function from those objects satisfying $\mathbb{N}$ to those satisfying set. More concretely, this amounts to the conjunction of the following three claims:

- $\mathcal{V}(V)$

- $\mathbb{N}, S$ satisfy $\mathrm{PA}_{\Diamond}$ (the categorical description of the numbers given in appendix F.3).
- ρ is a function from $\mathbb{N}$ to set

Note that we prove in lemma F.3.2 that it's logically possible to have $\mathbb{N}, S$ satisfy $\mathrm{PA}_{\Diamond}$.

I will often use the $\vec{V_a}$ notation to denote the pair V_a, ρ_a. And I will use $\Diamond_{\vec{V_n}}$ to abbreviate claims of the form $\Diamond_{\mathrm{set}_n, \in_n, \mathrm{ord}_n, @_n, \leq_n, \rho_n, \mathbb{N}, S}$ (and similarly for $\Box_{\vec{V_n}}$. Note that we use the same relations $\mathbb{N}, S$ for every $\vec{V_i}$.

We can now define a notion of extension for interpreted initial segment.

Definition 8.4.2 (Interpreted Initial Segment Extension)**.** The interpreted initial segment $\vec{V_b}$ extends $\vec{V_a}$ while assigning x written $\vec{V_a} \leq_x \vec{V_b}$ just if

- $V_a \leq V_b$
- $\mathscr{V}\left(\vec{V_a}\right) \wedge \mathscr{V}\left(\vec{V_b}\right)$
- $(\forall n \mid \mathbb{N}(n))\,(\rho_a(n) = \rho_b(n) \vee n = \ulcorner x \urcorner)$

My strategy will be to translate the set theoretic sentence $(\exists x)(\forall y)\,[x = y \vee \neg y \in x]$ with a potentialist claim about what is conditionally logically possible, given the structural facts about how some relations $\mathrm{set}_1, \in_1 \ldots \rho_1$ apply as follows.

8.5 The final product

With these definitions in place we can give a translation of $(\exists x)(\forall y)\,[x = y \vee \neg y \in x]$ as follows

$$\Diamond\left(\vec{\mathscr{V}}\,(()\,\vec{V_1}) \wedge \Box_{\vec{V_1}}\left[\vec{V_2} \geq_y \vec{V_1} \rightarrow \rho_2(\ulcorner x \urcorner) = \rho_2(\ulcorner y \urcorner) \vee \neg\rho_2(\ulcorner y \urcorner) \in_2 \rho_2(\ulcorner x \urcorner)\right]\right)$$

However I'll make one final change to the strategy illustrated above, to allow us to treat quantifiers in a uniform fashion. In the above examples the first quantifier had to be treated in a special manner as (the relations abbreviated by) $\vec{V_1}$ were not required to extend any $\vec{V_0}$. To this end, our translations will introduce a $\vec{V_0}$ and insist that $\vec{V_1}$ extend $\vec{V_0}$. Thus, for example, my official translation of $(\exists x)(\forall y)\,[x = y \vee \neg y \in x]$ is actually:

$$\Box\left(\vec{\mathscr{V}}\left(\vec{V_0}\right) \rightarrow \Diamond_{\vec{V_0}}(\vec{V_1} \geq_x \vec{V_0} \wedge \Box_{\vec{V_1}}\left[\vec{V_2} \geq_y \vec{V_1} \rightarrow \rho_2(\ulcorner x \urcorner) = \rho_2(\ulcorner y \urcorner) \vee \neg\rho_2(\ulcorner y \urcorner) \in_2 \rho_2(\ulcorner x \urcorner)\right]\right)$$

8.5.1 Recursive Definition of Potentialist Paraphrases

I will now describe recursive principles which let us translate every sentence in the first-order language of set theory into a claim about logically possible extendability.

First we define a partial paraphrase function t_n. Intuitively, $t_n(\phi)$ transforms a set theoretic formula ϕ into the a potentialist claim about the possible extendablity of the structure V_n where free variables are filled in by the assignment function ρ_n (coded by our assignment relation R_n).

Definition 8.5.1 (Potentialist Translation)**.** For any number n and set theoretic formula ϕ

- $t_n(x_i \in x_j) = \rho_n(\ulcorner x_i \urcorner) \in_n \rho_n(\ulcorner x_j \urcorner))$
- $t_n(x_i = x_j) = \rho_n(\ulcorner x_i \urcorner) = \rho(\ulcorner x_j \urcorner))$
- $t_n(\neg\phi) = \neg t_n(\phi)$
- $t_n(\phi \vee \psi) = t_n(\phi) \vee t_n(\psi)$
- $t_n(\phi \wedge \psi) = t_n(\phi) \wedge t_n(\psi)$
- $t_n((\forall x)\phi(x)) = \Box_{V_n}\left[\vec{V}_{n+1} \geq_{\mathbf{x}} \vec{V}_n \rightarrow t_{n+1}(\phi)\right].$
- $t_n((\exists x)\phi(x))$ is the claim that $\Diamond_{V_n}\left[\vec{V}_{n+1} \geq_{\mathbf{x}} \vec{V}_n \wedge t_{n+1}(\phi)\right]$

The translation of a set theoretic sentence ϕ is

$$t(\phi) = \Box\left[\mathcal{V}(V_0) \rightarrow t_0(\phi))\right]$$

In the above definition recall that $\Box_{\vec{V}_n}$ $(\Diamond_{\vec{V}_n})$ abbreviates a claim about what is logically necessary/possible holding fixed the facts about $\text{set}_n, \in_n$ $, \text{ord}_n, @_n, \leq_n, \mathbb{N}, S, \rho_n$

In what follows, I will, consistent with our general policy for functions, write $\phi(\rho_n(\ulcorner x_i \urcorner))$ to abbreviate claims of the form $(\exists k)\rho_n(\mathbf{i}, k) \wedge \phi(k)$. Moreover, since I will always subscript all the relations $\rho_n, \mathbb{N}, S$ whenever I subscript ρ_n I will assume that any $\Box$ or $\Diamond$ subscripting ρ_n actually subscripts $\rho_n, \mathbb{N}, S$.

8.5.2 Atomic Predicate Use Reducing Trick

If desired, in the above potentialist paraphrases we can replace V_j with $V_{j \bmod 2}$ without affecting the truth value of the translation. This allows us to translate sentences with arbitrarily many quantifier alternations using a fixed finite number of atomic relations.

For readability I won't employ this more economical paraphrase strategy in any of the arguments below. So I will argue as though we have access to a countable infinity of different atomic predicates, which we could use to translate sentences with quantifiers nested to arbitrary depth[2]

8.5.3 Equivalence of Approaches

It's useful to observe that my choice to add a base initial segment $\vec{V}_0$ into my translations is purely a matter of convenience as the two translation schemas turn out to be logically equivalent in my system.

So, for example, the straightforward paraphrase for $\exists x \phi(x)$ would be

$$\Diamond \left[(\mathscr{V}(\vec{V}_1) \wedge t_1(\phi(x))) \right]$$

and this is (in the formal system I propose) logically equivalent to my official paraphrase for $\exists x \phi(x)$

$$\Box(\mathscr{V}(V_0) \rightarrow \Diamond_{V_0} \left[\vec{V}_1 \geq \vec{V}_0 \wedge t_1(\phi) \right])$$

[2]However, I'd further suggest (but won't prove here) that all the formal *proofs* of ZFC axioms I show how to give below (not just all my potentialist paraphrases) can be done in a language of logical possibility with only finitely many atomic relations.

See lemmas H.4.1 and H.4.2 for a proof of this equivalence.

Chapter 9

How This Helps

In this chapter I will conclude Part II of this book, by discussing how the particular version of potentialist set theory we have developed lets us solve the problems noted for other versions of potentialism in Part I. I will also argue that it has some additional virtues and consider some objections.

9.1 Previously Mentioned Advantages

9.1.1 Avoiding Set Existence Awkwardness

Recall that in section 4 I raised a worry for Linnebo's 'Parsonian' potentialism arising from the question: what sets actually exist? Parsonian approaches to potentialist set theory cashed out set theoretic claims as modal claims about what sets could exist, flowing from (among other things) facts about the metaphysical nature and essence of sets, and how this constrained what sets could possibly exist.

As a result, Parsonian views seemed committed to acknowledging a fact about what sets (literally) actually exist, alongside all the modal facts about what sets could exist which they use to interpret mathematicians' talk of sets. Thus they seemed to have to posit to a kind of arbitrary fact about where the hierarchy of sets happens to stop (perhaps just as much as the actualist views). They also seemed committed to a kind of strange double meaning, whereby there are sets that exist but mathematicians' talk never refers to these, and only concerns itself with modal claims about what sets could exist. One might ask: in what sense are these abstract objects 'sets' if they are never what's referred to by mathematicians' set talk?

Like other Putnamian formulations of potentialism, my proposal avoids this family of worries. For it doesn't invoke a concept of set or the essence of sethood or (strictly speaking) what sets could exist. It only talks about 'what sets are logically possible' in the following sense. It talks about how it would be logically possible for some objects (considered under some relations) to form the kind of structure which actualists think standard width initial segments of the iterative hierarchy of sets must have. Thus, it isn't committed to a fact about what sets actually exist.

Specifically (following Putnam rather than Hellman), my particular potentialist translations of set theoretic claims take the form of claims about how it would be logically possible for non-mathematical predicates and relations of certain arities to apply. For example, one can write my proposed potentialist paraphrase of a given set theoretic sentence (say the Pairing Axiom in ZFC) as a claim about how it would be logically possible for there to be some penciled dots related by penciled arrows (as well as a number of other relations[1]). One could also write my proposed potentialist paraphrase of the same set theoretic sentence as a claim about how it would be logically possible for some flamingos and bears to be related to one another by admiration, fear and curiosity. (Happily, in the formal system I propose below, this pair of different translations for the same set theoretic sentences will be immediately logically equivalent to one another).

Perhaps allowing for this kind of free choice about which non-mathematical properties and relations to use will seem weird. One might ask: but which of the above of the formulations (invoking specific first order predicates and relations) should we say that set theorists' assertions really mean?

Here are two possible answers to that challenge, each of which I think is fairly satisfactory. First we might say that actual set theorists' talk is ambiguous between all these different logically equivalent formulations or even that there is no fact of the matter[2]. This position is analogous to a common position on Benacerraf's challenge to set theoretic reductions to say which specific collection of sets we should identify with the natural numbers (adopted by, e.g., Parsons in [72] and McGee in[67]). It is common to say that ordinary mathematical talk of numbers only has a definite reference up to isomorphism, being ambiguous between all the different copies of the natural number structure within the hierarchy of sets.

[1] My translations require relations for ord, $<$, @ as well as set and $\in$.

[2] One might even maintain the question itself is a kind of category error like asking who is 'the' witness to the truth of the claim 'There is a man who is more than six feet tall.'

Second, I could lean on the fact (noted in the introduction) that I'm only proposing that thinking of ordinary set theoretic utterances as shorthand for (something like) my potentialist paraphrases would be a useful *sharpening* of our current set theoretic concepts and practice. Given this, I'm not committed to saying that any version of my potentailist paraphrases is what set theorists currently mean. At most I'm committed to the claim that it would be a good idea to for set as the theorists to adopt (something like) my potentialist paraphrases.

9.1.2 Clarifying the Modal Notion at Issue

I suggested in chapter 4 that we face a worry about justifying our confidence in set theory, if we try to use Linnebo's potentialist paraphrases as part of a proposed foundation for set theory. I have suggested that Linnebo's notion of intepretational possibility isn't clear enough (or isn't determinate in the right ways) to motivate strong confidence in truth of many axioms about involving it. I have attempted to avoid this problem by using an independently motivated and intuitive notion of logical possibility in my potentialist paraphrases.

As noted above, one might try to avoid this problem by taking Linnebo's axioms to implicitly define the notion he is talking about. And indeed there are surely many interesting modal notions and variants on the concept of logical possibility we can study.

However this multiplicity of modal notions does not imply that any syntactic rules which we write down will correspond to a legitimate concept, or even be syntactically consistent. If we just write down arbitrary rules, we run the danger of adopting a modal version of Prior's [79] spurious logical connective 'tonk', which had the introduction rules for 'or' and the elimination rules for 'and' - and thereby allowed one to derive every sentence from every other sentence if accepted. So if we just think about Linnebo's modal principles as stipulative definitions, we can't justify confidence that they are even syntactically consistent.

In contrast, by providing an intuitive intended meaning for our modal notion $\Box$ and $\Diamond$ (via the notion of conditional logical possibility), which the axioms I will propose can be seen as intuitively true of, we banish concerns that these axioms are inconsistent. Above I argued that those attempting to provide a foundation for potentialist set theory must find a notion of possibility which we have strong enough independent intuitions about (and confidence in the cogency of our conception of) to justify ZFC axioms. And this is what I have tried to provide.

One might wonder if my notion of conditional logical possibility could be advanced as a way of understanding Linnebo's notion of interpretational possibility, and combined with his formal development of potentialist set theory. This would be nice, but here are some prima facie roadblocks for that project. First, Linnebo's notion of interpretational possibility is not supposed to be subject matter neutral (but only to characterize a sense in which more/different *sets* could exist). Second, (as we will see in the discussion of the Barcan Marcus formula below) Linnebo proposes axioms which are explicitly stated to rule out possibility of worlds smaller than the actual one, whereas it seems clearly logically possible that, e.g., only two things exist [3]. Thirdly, various principles which Linnebo takes to follow from the nature of sets (and uses in his development of potentialist set theory), don't seem to be logically necessary principles that apply to all objects.

9.2 Metaphysical Non-Specialness

A possible advantage of the potentialist set theory developed here concerns the question of how metaphysically special we should take the iterative hierarchy conception of sets to be.

One might argue that Linnebo and Parsons' potentialism is committed to the iterative hierarchy conception to be deeply natural/special (there are sets, which have an essence which requires exactly this), in a way that Putnamian versions of set theory can avoid.

For Putnamian potentialism has a somewhat if-thenist/modal structuralist flavor. It combines naturally with the idea that we just happen to have developed set theory by talking about how one logically possible kind of structure (an iterative hierarchy of sets) could exist and be extended, but we could equally well do mathematics by talking about the logical possibility and extendability of some other kind of structure.

[3] To be fair, one might argue that my notion also falls short of capturing the full intuitive notion of logical possibility, because axioms I will propose below let us derive that it is not logically possible for there to be a completely empty universe, while it is intuitively logically possible for no objects to exist. See the remarks about developing a version of my project which uses inclusive logic in the section on the Barcan Marcus formula below.

And arguably this feature of my/Putnam/Hellman's view provides an advantage over Parsonian theories with regard to metaphysical parsimony and access worries. For the Putnamian isn't committed to saying that there's something metaphysically special about the iterative hierarchy structure (as opposed to other ways of developing a set like concept which evades Russell's paradox) or the claim that set theorists have managed to latch on to such a uniquely special set-like concept.

Perhaps this laid back approach to the significance of the iterative hierarchy conception also fits better with mathematical practice. For (one might argue) the mathematical community just looked for *one* coherent conception set-like structure which avoided Russell's paradox and was generous enough to provide a foundation/larger mathematical universe within which other mathematical structures could be found and related to one another. They didn't do anything which looked like comparing different options for foundational structures which would do that work and inferring that the iterative hierarchy conception metaphysically superior.

9.3 Avoiding Second Order Logic

9.3.1 Nominalistic Qualms

Now let's compare the version of potentialism developed in this section with Hellman's potentialism for a moment. Aside from the concern about justifying potentialist use of ZFC axioms noted above (which will be answered by part III of the book), the main worry I considered was for Hellman's potetialism was his own concern that his use of second order logic generated undesirable ontological commitments; isn't second order quantification (as Quine suggested) just set theory in disguise?

Now this problem for nominalists is not really a problem for those (like myself) who are happy to allow the existence of various kinds of mathematical objects (like the natural numbers), but advocate potentialist set theory on the basis set theory speific concerns about the Burali-Forti paradox. For (as discussed in section 2.5.2) no version of the Burali-Forti paradox seems to arise for talking about a single layer of classes over the first order objects. Thus the motives for denying that there's an actualist hierarchy of sets don't apply to rejecting second order comprehension.

However I think adopting my proposed form of potentialism can help philosophers who do want to be nominalists solve this problem by eliminating all appeals to second order logic, in favor of a notion of conditional possibility which (as I argue in [5]) is very close to notions Hellman already accepts the nominalistic credentials of and uses to make sense of applied math.

9.3.2 Conceptual Simplification

Quite aside from concerns about nomainalism, one might argue that my potentialist set theory's avoidance of second order quantification provides an has an advantage on the front of conceptual parsimony.

Hellman uses: first order logical vocabulary, second order quantification, logical possibility and (to make sense of applied mathematics, which I haven't discussed here) something like an actuality operator.

Linnebo uses: first order logical vocabulary, plural quantification, interpretational possibility, and the notions of set and elementhood.

My paraphrases use: first order logical vocabulary and the conditional logical possibility operator.

Insofar as a single modal notion turns out to be able to do both the following jobs in potentialist translation, it is appealing to use it.

- describe the (first order logic transcendent) structure of the partial hierarchies of sets potentialists want to consider
- articulate the sense in which some such structures are possible and could be extended in certain ways.

To put this idea about conceptual parsimony another way, there can seem to be a kind of inelegant conceptual double counting in employing both the $\Diamond$ of logical possibility and a notion of second order quantification while treating these as unrelated primitives (as Hellman does). For, intuitively, there's something in common between the way we consider 'all possibilities' for how some first order predicates could apply when evaluating logical possibility and the way we consider 'all possibilities' for choosing some first order objects from a given collection when considering what second order objects exist.

Perhaps actualists about set theory can straightforwardly explain this similarity. For they can define both notions in terms of what sets exist[4]. In particular they can appeal to the same notion of 'all subsets over a given first order domain' when defining logical possibility in terms of the existence of a model and when cashing out second order quantification in terms of sets existence.

But potentialists cannot do the same. For we potentialists understand set existence in terms of logical/interperetational/whatever possibility, rather than the other way around. So we cant account for the sense of conceptual overlap between the notions of logical possibility and second order quantification by cashing out both notions in terms of set theory. Thus we lose the above benefit and, e.g., Hellman and Linnebo wind up treating logical possibility and second order quantification as just separate conceptual primitives.

Happily, however, we can solve this problem if we embrace the notion of conditional logical possibility, for this single notion a can be used to articulate and analyze both claims about logical possibility and second order quantification[5][6]

[4]Or at least they can do this if we bracket Field's objection to identifying claims about logical possibility with claims about set theory discussed in 6.

[5]We have seen how to do this in for the purposes of second order claims needed to formulate potentialist set theory. See REDACTED for an argument that we can reformulate second order claims more generally

[6] On the topic of second order logic, one might worry that my favored potentialist paraphrase strategy above is worse than Hellman's insofar as I have only provided paraphrases for sentences of first order set theory, and not paraphrases for second order statements which attempt to quantify over the hierarchy of sets.

I this isn't much of a cost, since what Hellman says isn't very attractive. For example, in section 2.3 or *Mathematics Without Numbers* himself notes that his potentialist translation of second order replacement is unsatisfying insofar as it doesn't imply (his potentialist translation of) first order replacement.

Also I think the core potentialist idea suggests that second order quantification over the

9.4 Quantifying-In

Finally I will argue that the fact that my formulation of potentialist set theory avoids quantifying into the $\Diamond$ of logical possibility (and indeed my language of logical possibility does dot allow any quantifying in) has some philosophical benefits.

9.4.1 Sidestepping Controversy

In Part I we saw that the major existing formulations potentialst set theory employ quantifying in to the $\Diamond$ of logical possibility. That is, they use sentences like $\exists x \Diamond R(x)$, where the logical possibility operator is applied to a formula with free variables.

However there are significant controversies about the truth conditions (and indeed meaningfulness) for very simple sentences involving quantifying, which can make it difficult to provide uncontroversial foundations for versions of potentialist set theory which employ quantifying in.

For one thing, there is disagreement about whether all pairs of things that are actually distinct are necessarily distinct[7].

There is also disagreement about what to say about statements which quan-

full hierarchy of sets *shouldn't* make sense. There isn't some fixed first order collection, the sets, such that we can ask what subsets of it exist. But for those who are not convinced, I should note that my system can also mimic Hellman's translations.

[7] One might try to argue that it's metaphysically and hence logically possible for there to be two people who could have been one person as follows. Suppose that two people are formed by a contingent event of person splitting. One might think these people are distinct but could have been identical.

tify into a world where an object doesn't exist. For example, Kripke's[54] approach (which Hellman endorses) allows sentences like $(\exists x)\Diamond[\neg\mathrm{Fox}(x) \wedge (\forall y)\mathrm{Fox}(y)]$ to be true, a consequence which Williamson[108] and others have argued is extremely counterintuitive[8]. These controversies can raise doubts about whether our intuitions about quantifying in are reliable while, to my knowledge, no analogous paradoxes arise in the system I lay out[9] (and the absence of quantifying in makes it hard to see how they could).

There is also a Quinean strand of argument which claims that quantifying into modal contexts is meaningless. I take Quine's problem with quantifying in, in [84], to be that he dislikes the "Aristotelian essentialism" of saying that some properties to belong to an object like the number 7 essentially (e.g., being less than 9) while others apply only contingently (e.g., being the number of planets). After all, taking there to be such an abundance of facts about essences can seem arbitrary. My language of logical possibility eschews cross-logically-possible-world object identification of any kind (e.g., cross-world equality or counterpart relations) as well as quantifying in[10]. Thus and worries like Quine's don't even get off the ground. And it seems, at least, rhetorically desirable to demonstrate that potentialist set theory

[8] While this debate is commonly conducted in terms of metaphysical possibility, it naturally raises similar concerns for logical possibility.

[9] My proposed account of set theory is compatible with taking Williamson to show that any modal notion *which allows quantifying in* (such as metaphysical possibility) must have a fixed domain – provided one thinks it doesn't make sense to quantify in to logical possibility. Of course, it's not compatible with taking Williamson to show that every modal notion must have a fixed domain.

[10] For, recall that when evaluating conditional logical possibility we consider logically possible scenarios which preserve structural facts of how some relations $R_1, \ldots, R_n$ apply, rather than facts about which objects they apply to. So advocates of my version of potentialist set theory aren't thereby committed to taking essence claims about e.g. what it would be logically possible for a given object to do to be meaningful.

doesn't require quantifying in or similarly controversial notions.

So much for disagreements about the meaningfulness of quantifying in and the truthvalue of basic statements involving quantifying in. Now, in itself, the existence of such controversy isn't necessarily a problem for any particular formulation of set theory which uses quantifying in, like Linnebo's or Hellman's. For perhaps one of these philosophers is getting the right answer regarding this disputed question, or perhaps there are different equally legitimate things the syntax of quantifying into the $\Diamond$ can mean, and each theorist is getting things right regarding the semi-formal language he is using.

However the existence of such controversies does raise a problem for using formulations of potentialist set theory involving quantifying in (like Linnebo's and Hellman's) in the foundational project I'm attempting. For my aim is to indicate a clear modal notion and then show that we can justify the potentialist translations of the ZFC axioms from modal principles which seem clearly and uncontroversially true. By doing this we can vindicate mathematicians' use of these axioms from a potentialist perspective, answer philosophical questions about the justification of replacement etc. But we can only hope to provide the kind of confidence and deduction from the obvious and intuitive one would intuitively like from a foundations for mathematics by latching on to uncontroversial principles.

Accordingly I think it will help me achieve the foundationalist aims of this book to formulate a language of logical possibility which doesn't allow quantifying in (and hence doesn't let one formulate the sentences whose truth-value/meaningfulness is so widely disputed), and show that all the modal claims we need to articulate and justify poentialist set theory can be stated within this language. I think that, rightly or wrongly, people are more disposed to agree about what's conditionally logically possible (once the meaning of that concept has been roughly gestured at) than what's de re logically possible for individual objects (perhaps because of the felt connection between the latter facts and questions about de re metaphysical possibility). And given this fact, it's pragmatically useful for someone attempting to provide uncontroversial modal axioms, and hence an uncontroversial foundation for set theory, to use the former notion and not the latter.

9.4.2 Shyness and Compossibility Worries

By avoiding quantifying in, my proposal also avoids a different worry, about 'metaphysical shyness' which Linnebo raises for Hellman in [61]. Linnebo suggests that to make sense of set theory via Hellman's proposal (which takes potentialist set theory to reflect general constraints on the behavior of all objects) we will want to say (so to speak) any two logically possible structures could be combined, yet it seems that there could be metaphysically shy objects such that the essence of these objects makes it metaphysically (and even logically?) impossible for them to exist side by side.

Linnebo formulates a kind of shyness worry about metaphysical possibility, writing "Do we really know that there cannot be 'metaphysically shy objects, which can live comfortably in universes of small infinite cardinalities, but which would rather go out of existence than to cohabit with a larger infinite number of objects? Another variant of this problem arises when we consider applied set theory. In order to apply set theory to some given objects, the approach in question presumably needs it to be possible for these objects to coexist with strongly inaccessibly many objects, arranged so as to make up a standard model built on top of these ur-elements. But this claim will be false if there can be objects that necessarily cannot coexist with inaccessibly many other objects."[61]. Then he suggests that a similar shyness worry can be raised for logical possibility. He also suggests that logical possibility allows for an analog of the problem of metaphysically incompatible objects (like two metaphysically possible knives formed by joining a single handle with different blades).

I take my formulation of potentialist set theory in terms of conditional logical possibility to avoid this problem of metaphysically shy objects, insofar as it avoids all quantifying in to the $\Diamond$ of logical possibility. My approach doesn't make any claim about the essences of objects, just claims about what is possible while preserving structural facts about how some properties apply (something that does not require preserving the specific objects which these properties apply to).

9.4.3 Barcan Marcus

A final worry about quantifying in concerns the converse Barcan Marcus formula $\Box(\forall x)\phi(x) \rightarrow (\forall x)\Box\phi(x)$ and the question of whether everything exists necessarily. Note that the converse Barcan-Marcus formula is provable in most (reasonably strong) quantified modal logics and easily allow one to derive the conclusion that everything exists necessarily. In particular, if we take $\phi(x)$ to be $x = x \to (\exists y)(y = x)$ we easily see that $(\forall x)\phi(x)$ is logically true and thus infer $(\forall x)\Box\phi(x)$, i.e.,

$$(\forall x)\Box x = x \to (\exists y)(y = x)$$

Philosophers who use quantifying in to handle potentalist set theory can respond to this problem in different ways. Linnebo handles this problem by embracing it, and saying that the relevant modal notion of 'possible construction' relevant to set theory doesn't allow for the possibility of smaller worlds in [60].

I will (in effect) handle this (in largely the spirit that Hellman and Kripke [45][55] do) by saying that the necessitation rule only applies to complete sentences. And perhaps this is intuitively motivated; we wouldn't say it's logically necessary/a tautology that $x = x$, because formulas with free variables aren't even sentences so they don't have truth values, and hence can't be necessarily/logically true.

More specifically, we have seen that the language of logical possibility doesn't treat such sentences as meaningful. And I will provide a formal system for reasoning within this language whose individual principles and inference rules are easy for us to evaluate because it (in effect) cleaves good reasoning about logical possibility up into two parts.

- In one part we use standard first-order logic to reason about a given logically possible scenario/what an arbitrary logically possible scenario must be like.

- In another part we use special modal-structural principles to establish which scenarios are logically possible, and 'transfer' facts about one scenario to another.

An alternative approach would be to allow quantifying in, but use an inclusive logic[70] . Switching to an inclusive logic would let us block the above argument for the Barcan-Marcus formula by blocking the initial proof that '$(\exists y)(y = x)$, rather than the application of necessitation to this formula in the last sentence as inclusive logics neither assume that all singular terms refer to members of the domain nor that the domain is non-empty.

I think this strategy is prima facie quite appealing, because it would allow us to capture the intuitive logical possibility of entirely empty domains. However because, as a matter of sociological fact, no inclusive logic is currently widely accepted, I have preferred to sacrifice intuitions about empty domains and use classical first order logic rather than arguing for new views on both first order logic and set theory in this book. I leave the task of reworking the formal system and proofs provided in this book to fit with any inclusive logic that may gain wide acceptance in the future.

Chapter 10

Naturalist Worries

Above I've argued for a potentialist approach to set theory, on which set theoretic talk is best 'explicated' as making modal claims about how it would be *possible* to have various hierarchy-of-sets like structures extending one another. This potentialist approach to set theory counts as a nominalist, insofar my logical regimentations for pure set theoretic claims don't quantify over mathematical objects.

This approach to set theory might raise worries for philosophers who embrace extreme naturalism of a kind that is currently rather popular in philosophy of mathematics[63] [14]. Such philosophers will wonder: is my choice of these nominalistic regimentations adequately mathematically/scientifically motivated or does it have only philosophical motivations (something the extreme naturalist would consider inadequate)?

In this chapter, I will respond to a version of Burgess and Rosen's influential dilemma for mathematical nominalists, which nicely sharpens the above worry. In a nutshell, Burgess and Rosen argue that nominalist logical regimentations of scientific sentences are inadequately naturalistically motivated, whether we consider them as claims about what mathematicians and scientists *should* say (because they wouldn't be published in relevant scientific journals) or as claims about what scientists currently *do* say (because they lack adequate motivation from linguistics and cognitive science).

I will argue that my potentialist logical regimentation of set theory may be adequately naturalistically motivated, even if the above claim about publication in core mathematics journals and linguistics/cognitive science data are correct.

10.1 Burgess and Rosen

In [14] Burgess and Rosen propose the following influential argument-by-dilemma that nominalistic logical regementations of pure and applied mathematics lack adequate naturalistic motivation.

Consider some proposed nominalistic logical regimentation of a scientific theory that appears to quantify over mathematical objects. This proposal must be intended either as a 'hermeneutic' proposal, clarifying what people currently mean by the theory or a 'revolutionary' proposal concerning what they should mean. But (they argue) typical convoluted nominalistic paraphrases of applied mathematical statements in the sciences would seem to be

- too psychologically and linguistically unmotivated to be a plausible hermeneutic theory of what people currently mean.
- too unmotivated by the standards of the scientific discipline in question to be a plausible revolutionary theory of what scientists should say/mean. For example, nominalistic regimentations of a physical theory would generally not be accepted by physics journals.

Thus it would seem that, in either case, this logical regimentation (and the theory it is used to defend) should probably be rejected.

This general argument against nominalism has prima facie force against my limited nominalism about set theory as well. For one might argue that my potentialist formulations of set theory are implausible as either revolutionary or hermeneutic proposals for the following reason.

- It's unclear whether mathematics journals would publish work in potentialist set theory, and core mathematics journals probably would not[1].
- Linguists, cognitive scientists or grammarians studying set theoretic talk in ordinary contexts (including publication in core mathematics journals) would probably give Platonist logical regimentations not potentalist ones.

so I will now try to answer this worry.

[1]On the other hand, journals on the boundaries of mathematics like the Review of Symbolic Logic certainly do publish work on potentialist set theory (c.f. Linnebo and Hamkins' "The modal logic of set-theoretic potentialism and the potentialist maximality principles[42]) and mathematicians concerned with the questions about adding new axioms to set theory like Hamkins take an interest in it.

10.2 Naturalism Rejecting Response

One obvious response to the argument above is simply to reject the naturalistic presuppositions that adequate motivation for adopting a theory must be mathematical or scientific in character. An untroubled friend of traditional philosophical metaphysics can happily and frankly advocate the 'revolutionary' line that their nominalistic regimentation reflects what we should say when speaking seriously and literally (in a regimented language which displays our commitments as per Quine's criterion [86]), as we do when discussing metaphysics.

They can say that we have good *philosophical* reasons for preferring such paraphrases, which are to be taken just as seriously as mathematical and scientific ones[2]. If (as Burgess and Rosen suggest) these nominalistic regimentations of physical theories are not what we should say when submitting to physics journals, this may simply be because articles in physics journals are (for obvious practical and division of labor reasons) not attempting to speak completely literal in a logically regimented language which exposes the metaphysical structure of reality when we apply Quine's criterion. Articles in science journals are instead written in unregimented natural language

[2]In *A Subject with no Object* Burgess and Rosen also argue individually against various traditional philosophical motivations for nominalism about all mathematical objects: access worries that accepting ontological realism makes human accuracy about mathematics mysterious, appeals to Occam's razor, and resistance to necessary, abstract etc. objects.

I don't take this first cluster of worries to be a problem for my potentialist set theory, insofar as I have suggested a subject matter specific motivation (the Burali-Forti paradox) for preferring a nominalist explication of set theory: Adopting potentialist set theory lets us articulate a conception of the height of the hierarchy of sets (in terms that we seem to understand even after thinking about Burali-Forti, and reconcile this with intuitive logical-combinatorial facts about all actual structures being extendable.

which is easier to work with – and perhaps helpfully lets one bracket certain metaphysical questions[3].

However, some readers may feel more sympathetic to the naturalist intuitions invoked by Burgess and Rosen's dilemma above.

10.3 Naturalistic Response

So I will now argue that the argument above fails, even if one accepts the demand for naturalistic motivation. Suppose the assumptions about mathematical and scientific journals and motivation from cognitive science and linguistics above are correct. I will argue that this doesn't suggest that there isn't adequate mathematical motivation for adopting potentialist set theory for the following reason.

Providing mathematical foundations for mathematical subdisciplines (which aren't expected to correspond to the surface grammar of what we'd say in journals devoted to these subdisciplines) is already an accepted and apparently fruitful part of contemporary mathematical practice. Mathematicians already draw a distinction between what it's right to say in normal contexts (including the classroom or and mainstream mathematical journals) and what it's right to say in certain unusually pedantic contexts of foundational investigation.

Take, for example, the classic set theoretic foundations for analysis invoked

[3] I have in mind questions like whether there's an abstract object 'electronhood' or merely a property? For example, writing up a physical theory in a logically regimented language which Quine's criterion can be applied might require one take a stand on this.

above. There is plausibly no linguistic/cognitive sciences motivation for logically formalizing the assertions which are or should be in an analysis journal as quantifying over sets and talking about Dedekind cuts rather than merely talking about real numbers and not asserting that they have any further structure. Success at a providing set theoretic foundation for analysis (as motivated by the need to solve problems and paradoxes within analysis) doesn't require providing an analysis which is motivated by looking at the surface grammatical form of what one does or should say in most contexts.

Similarly, I propose that success at providing a modal foundational for set theory (as motivated by the need to solve problems within set theory like the Burali Forti paradox) doesn't require providing logically regimented sentences which simply match the surface grammatical form of what people do or should will say about set theory in most contexts. For contemporary mathematical practice itself seems to clearly acknowledge/allow that there can be good mathematical reason for adopting logical regimentations for mathematical talk which are far more elaborate than what we do or should say in most mathematical contexts.

We might think of these foundational proposals as accounts of what one should say in a context with the following features. One has plenty of time (so there is no need for abbreviation) and no need to teach others (so there's no need for technically false simplifications). But one lays oneself open to arbitrary pedantic or strange questions, e.g., questions which connect very 'disparate' parts of one's web of mathematical beliefs. And one tries to apply ones concepts to questions of types not generally considered (e.g., taking limits of certain strange functions which are not physically natural).

Theoretically, it also makes sense that mathematicians would make such a distinction between what should be said in foundational/reductive contexts vs. in the course of something like Kuhnian normal science where we know how to get right answers by employing familiar ways of talking and techniques. For, on the one hand, it it is useful for them to precisify their terms if they are reasoning at the edges of normal practice, in cases where paradox threatens or its desirable to apply concepts from one domain to new areas. But, on the other hand is also desirable for mathematicians and scientists to continue with an apparently working practice of studying something and not to commit themselves to any specific foundational analysis of what is going on 'under the hood'. Journals devoted to an area of normal science which seems to be going well, needn't bother attempting to further analyze their terms. And perhaps articles written within such patches of functioning normal science shouldn't take the risk of committing themselves to one answer to foundational questions over another.

If this division between normal and foundational mathematical contexts is accepted, then we should not be troubled by the fact (if it is a fact) that most mainstream mathematics journals wouldn't want to publish some theory which is proposed as a foundation for mathematical subdiscipline. So, I claim that even if the assumptions about publication in mathematical journals (and the results of linguistically analyzing ordinary set theoretic talk) are correct, naturalistic philosophers can take the account of potentialist set theory developed here seriously as a story about what we should say in the special pedantic context of foundational debate.

If one further says (as seems plausible, but optional, for the naturalist) that ones ontological commitments are given by applying Quine's criterion ('to be is to be the value of a bound variable') to ones best logically regimented theory of the world *as one would state it in pedantic/foundational contexts*[4], this amounts to a defense of nominalism about set theory against Burgess and Rosen's dilemma.

[4]This seems like a fairly reasonable way to proceed for a philosopher strong naturalist sympathies who wants to take ontological questions seriously at all. If someone says that they don't really believe in average plumbers and can cash out all their talk of such forms in other terms, it seems only reasonable to acquit them of ontological commitment to average plumbers (i.e., to read their ontology off what they say in pedantic/foundational contexts not everyday one).

However one might argue that very strong naturalist sympathies of the kind which drive Burgess and Rosen's dilemma (philosophical motivations count for nothing) should lead one to reject questions of traditional ontology all together. For example it's worth noting that *platonist* logical regimentations of antecedently well known physical theories likely also wouldn't be accepted by physics journals. Arguably the very idea of providing a logical regimentation of a theory and then applying Quine's criterion to it to say what exists depends on accepting a philosophically motivated project of ontology. So we shouldn't think of the 'science journals wouldn't accept it' point as motivating realism about mathematical objects over nominalism, but rather as suggesting a kind of quietism/agnosticism about all philosophical questions which can't be studied by the accepted methodology of some sciences, and hence rejecting both nominalism and platonism about mathematical objects alike.

Part III

Chapter 11

Content Restriction

In this final section of the book (Part III) we will develop a formal theory of how to reason about (conditional) logical possibility, and then use it to justify acceptance of the ZFC axioms. Chapters 12 and 14 propose some intuitively appealing axioms and inference rules for reasoning about conditional logical possibility. But first I must introduce content restriction, a key concept we will need to state these axioms and rules.

11.1 A motivating example

To motivate the idea of content restriction, consider purely number theoretic statements, i.e., statements whose syntactic form makes it clear that they only make a claim about the structure of the natural numbers, rather than a claim whose truthvalue might reflect the behavior of some larger universe of objects. Intuitively, the truth value of such a statement is completely determined by the structure of the natural numbers. That is, their truth values are completely determined by structural facts about how the relations $\mathbb{N}, S, +$ and $\times$ (S for 'successor') apply (call this the $\langle \mathbb{N}, S, +, \times \rangle$ structure) and don't depend on what other objects may or may not exist.

So, for example, consider the number theoretic claim that there are infinitely many twin primes. We know that if the current state of the world makes this statement true, then facts about the natural number structure alone suffice to do this. Because it only quantifies over the natural numbers, and only concerns itself with how the relations successor, plus and times apply to the natural numbers, we can't change its truth-value by adding or subtracting objects from the universe outside the extension of $\mathbb{N}$ or by tinkering with the extension of properties and relations other than $S, +$ and $\times$ (like 'is democratically governed' or 'is a spaceship').

Accordingly, if a purely number theoretic statement is true, we expect it to remain true under any changes to the world which don't effect the natural number structure. As time passes, empires may fall, the population of humans may rise and the total size of the universe may change radically. But, so long as the natural number structure remains fixed, the truth-value of purely number theoretic statements cannot change.

And that's not all. We further expect that the truth-value of any purely number-theoretic statement cannot be changed by the application of any conditional logical possibility operator which holds fixed the natural number structure (i.e., the facts about how $\langle \mathbb{N}, S, +, \times \rangle$ apply). For, intuitively, modifying what other objects exist outside the structure of the natural numbers which this statement quantifies over (and/or changing the extension of other relations which it doesn't employ) can't affect its truth-value.

This idea (that truth value of all purely number theoretic claims must be the same in all logically possible worlds which preserve the structure of the natural numbers) fits well with common intuitions about the significance of non-elementary proofs.

Thinking about how the natural numbers would relate to other larger mathematical structures, like the complex numbers, can be epistemicly helpful in discovering the answer to some purely number theoretic statements. Proofs of this kind are called non-elementary proofs. But we don't think about the existence of the complex numbers as helping make these number-theoretic statements true or false. Rather we think that, if true, the relevant number-theoretic statements must have been true 'all along', just because of what the natural numbers are like, and considering how the natural numbers are (or could be) related to the complex number just helps us see this fact.

Accordingly, it doesn't matter to our acceptance of a nonelementary proof whether we think the complex numbers actually exist, or merely that it would be logically coherent for (an instance of) the actual natural number structure to exist inside (an instance of) the complex numbers. Showing the twin prime conjecture is true and merely showing that it would have to be true in the relevant logically coherent scenario both suffice to establish its truth in the actual world.]

11.2 Generalizing this idea

Let us now generalize the above idea that the syntactic form of certain sentences ensures they are 'purely about the natural numbers structure' – and hence must remain true in all logically possible scenarios which hold facts about the natural number structure fixed. That is, such statements must remain true in all scenarios which hold fixed how the relations $\langle \mathbb{N}, S, +, \times \rangle$ apply.

So consider what we want to say about the natural numbers. Let ϕ be a statement whose syntax insures that it is purely about the natural numbers in the following sense:

- All quantifiers in ϕ are restricted to the objects satisfying $\mathbb{N}$. Thus (we can write it to make this explicit, i.e., so that) it only contains universal quantifiers as part of expressions of the form $\forall x(\mathbb{N}(x) \rightarrow \psi)$ and existential quantifiers as part of expressions of the form $\exists x(\mathbb{N}(x) \wedge \psi)$
- ϕ is a sentence in the language of numertheory, so it only contains relations on this list: $\mathbb{N}, S, +, \times$

Then (if we accept intuitions I've tried to pump above) we expect that ϕ cannot change in any conditionally logically possible scenarios which hold the facts about the natural numbers fixed. Accordingly, ϕ is actually true iff it is conditionally logically possible - holding fixed the natural number structure- that ϕ be true.

$\phi \leftrightarrow \Diamond_{\mathbb{N},S}\phi$

I will generalize this idea by considering other lists of relations $R_1...R_n$ (rather than just $\mathbb{N}, S$). I will define a syntactic property of sentences which intuitively ensures that a sentence is completely about (structural facts concerning) how some list of relations $R_1...R_n$ apply, so that its truthvalue (intuitively) must remain fixed in all conditionally logically possible scenarios which hold these relations fixed. I will call this property 'explicit content restriction'. And we will see that when a sentence ϕ is content restricted to some list of relations $R_1...R_n$ it is intuitively clear that $\phi \leftrightarrow \Diamond_{R_1...R_n}\phi$

However one little wrinkle arises in performing this generalization. In the case of the natural numbers, we thought of the structure of the natural numbers under the relations $S, +, \times$. And we said that purely number theoretic statements had their quantifiers restricted to objects in the extension of $\mathbb{N}$. But now we want to generalize this idea of 'only talking about' the structure determined by how some arbitrary list of relations $R_1...R_n$ apply. And this list of relations doesn't have one particular property distinguished as representing the domain.

So what exactly should it mean to talk about'the $R_1...R_n$ structure'? Specifically, what domain of objects should we consider the behavior of, under the relations $R_1...R_n$? I will handle this problem by (in effect) considering the domain of objects which *any one* of the relations $R_1...R_n$ apply to, under the relations $R_1...R_n$. So, for example, if our list of relations $R_1...R_n$ is cat(),loves() then an object is satisfies $Ext(R1...Rn)$ iff it is *either* a cat or a lover or a beloved. And we will consider the structure of objects determined by the relations $R_1...Rn$ (analogous to the natural number structure, in the original case) to be the structure of the objects in Ext(Cat(),Loves()) under the relations cat() and loves().

With just a little abuse of notation, we can call this domain of objects associated with the $R_1...R_n$ structure the *extension* of the list of relations $R_1...Rn$, and giving the formal definition below.

Definition 11.2.1 (Definition of Ext)**.** Let $\mathrm{Ext}(R_1, \ldots R_n)(y)$ abbreviate the formula

$$\bigvee_{\substack{1\leq i\leq n \\ 1\leq j\leq l_i}} (\exists x_1)\ldots,(\exists x_{j-1}),(\exists x_{j+1}),\ldots,(\exists x_{l_i})R_i(x_1,\ldots,x_{j-1},y,x_{j+1},\ldots,x_{l_i})$$

where l_i is the arity of R_i and $\bigvee_{\substack{1\leq i\leq n \\ 1\leq j\leq l_i}} \phi_{i,j}$ indicates the disjunction $\phi_{i,j}$ over all indicated values for i and j. Thus, $\mathrm{Ext}(R_1, \ldots R_n)(y)$ is the formula asserting that some tuple $\vec{v}$ including y satisfies some $R_i(\vec{v})$.

So I will take 'the $R_1...R_n$ structure to be the structure of objects in $\mathrm{Ext}(R_1, \ldots R_n)$ under the relations $R_1...R_n$. And I will try to define a syntactic property of explicit content restriction, such that the fact that a sentence ϕ is explicitly content restricted to the relations $R_1...R_n$ intuitively insures that ϕ only talks about this structure, so that its truth-value will be preserved in all conditionally logically possible scenarios which hold this structure fixed.

11.3 Formal Definition

So *explicitly content restricted* sentences are supposed have a syntactic structure which ensures that their truth-value is completely determined by the $\mathcal{L}$ structure (for $\mathcal{L}$ a certain list of relations), with the result that $\phi \leftrightarrow \Diamond_{\mathcal{L}}\phi$.

To motivate my definition, consider two examples.

The truthvalue of the sentence 'Every lover is loved by someone' $\forall x(\exists y\text{Loves}(xy) \rightarrow \exists z\text{Loves}(zx))$ is completely determined by facts about the Loves() structure in the sense defined above. That is, it only makes a claim about the structure of objects which are either lovers or beloveds, under the relation Loves.

We can see this by noting that it is logically equivalent to a sentence with quantifiers explicitly restricted to objects in Ext(Loves) and the considerations above apply.

$(\forall x \mid Ext(L)(x))((\exists y \mid Ext(L)(y))Lxy \rightarrow (\exists z \mid Ext(L)(z))Lzx)$

And we clearly cannot change the truth-value of the resulting sentence by

- adding or subtracting objects which neither person() nor loves() applies to from the universe
- changing the extension of predicates other than person() and loves()

So we can't change the truthvalue of this sentence without changing the Loves structure (in the sense defined above). Thus we intuitively have $\forall x(\exists yLxy \rightarrow \exists zLzx) \leftrightarrow \Diamond_L\forall x(\exists yLxy \rightarrow \exists zLzx) \leftrightarrow \Box_L\forall x(\exists yLxy \rightarrow \exists zLzx)$

In contrast, the truthvalue of the sentence 'Everything loves something', i.e.,

$(\forall x)(\exists y)(\text{Loves}(x, y))$, is not completely determined by the Loves() structure. For the existence of objects outside of this structure can make a difference to its truthvalue. Specifically it is logically possible that this sentence be true. But, given any world where this sentence is true, we can imagine a logically possible scenario which holds fixed the structural facts about how loves() applies at this world, but makes this sentence false by containing an additional object which loves does not apply to. Thus the truth-value of this sentence is not completely determined by the world's *Loves* structure. We might have $(\forall x)(\exists y)(\text{Loves}(x, y))$ but $\Diamond_{Cat,Loves}\neg(\forall x)(\exists y)(\text{Loves}(x, y))$

Roughly speaking, I will say that a sentence ϕ is explicitly content restricted to a finite set (note the notions of sets aren't presumed in the object language merely used to describe the meta-language and can be easily be eliminated[1]) of relations $\mathcal{L}$ iff only the relations from $\mathcal{L}$ are used in ϕ and every quantifier is restricted to range over elements that belong to some tuple in the extension of a relation in $\mathcal{L}$. [2] The definition below expresses this idea.

Note that I will frequently drop the braces and union symbols when talking about sets of relations. For instance, I will abbreviate the claim that ψ is content restricted to $\mathcal{L} \cup \{R\} \cup \mathcal{L}'$ simply as ψ is content restricted to $\mathcal{L}, R, \mathcal{L}'$ and write $\mathcal{L} = R_1, \ldots, R_n$ rather than $\mathcal{L} = \{R_1, \ldots, R_n\}$.

Definition 11.3.1 (Content Restriction)**.** A sentence ϕ is **explicitly content-**

[1] Sets merely serve as a convenient way to talk about syntactic properties of formulas in the language of logical possibility. Since we only make use of finite sets there is no need for a set-theoretic meta-language (we could directly give a computable enumeration of allowed inferences).

[2] The intuitive conception outlined above only makes sense for sentences, but it will define a notion of content restricted for formulas as well so we can keep track of whether the sentences built from them are content restricted.

restricted to a list $\mathcal{L}$ if it is a member of the smallest set S satisfying:

- $\bot$ is in S
- If v_i, v_j are variables the formula $v_i = v_j$ is in S
- If v_i is a variable and $R_i \in \mathcal{L}$ then $R_i(v_j)$ is in S
- If $\psi \in S$ and $\rho \in S$ then $\neg\psi$, $\psi \vee \rho$, $\psi \wedge \rho$ and $\psi \rightarrow \rho$ are all in S
- If $\psi \in S$ and $\mathcal{L}$ is non-empty then $\exists y(y \in \mathrm{Ext}(\mathcal{L}) \wedge \psi)$ is in S
- If $\psi \in S$ and $\mathcal{L}$ is non-empty then $\forall y(y \in \mathrm{Ext}(\mathcal{L}) \rightarrow \psi)$ is in S (note this case is added only for illustrative purposes as technically $\forall$ is merely an abbreviation for $\neg\exists\neg$).
- If $\phi = \Diamond_{\mathcal{L}'}\psi$, where ψ is a sentence and $\mathcal{L}' \subseteq \mathcal{L}$ then $\phi \in S$. Note that *ψ need not be in S*

Clauses 2-6 above express the basic idea from above: that if a sentence employs only quantifiers which are restricted to $\mathrm{Ext}(\mathcal{L})$ and relations in $\mathcal{L}$, then it makes a claim which is purely about the $\mathcal{L}$ structure, and its truthvalue must be completely determined by this structure.

Clauses 1 and 7 liberalize this definition slightly, by allowing two other basic ingredients to figure in sentences which are content restricted to $\mathcal{L}$.

Clause 1 allows sentences content which are content restricted to $\mathcal{L}$ to employ the logically false proposition $\bot$. This is motivated by the fact that $\bot$ is intuitively content restricted to any list of relations $\mathcal{L}$. Because there is no way to change $\bot$'s truthvalue at all, there is no way to change it while holding fixed the facts about the $\mathcal{L}$ structure.

Clause 7 allows sentences which are content restricted to $\mathcal{L}$ to employ claims about conditional logically possible given the facts about the $\mathcal{L}$ structure (or some part of it). Recall that $\Diamond_{\mathcal{L}'}\psi$ says that it's logically possible for ψ to be true, while holding fixed the structural facts about how the relations in $\mathcal{L}'$ apply. Accordingly only structural facts about the relations in $\mathcal{L}'$ should be able to make a difference to its truthvalue (remember that we don't allow quantifying into the $\Diamond$, so ψ cannot contain any free variables assignments of variables won't matter). So when $\mathcal{L}'$ is a subset of $\mathcal{L}$, we can't change the truth value of this sentence without changing how some relation in $\mathcal{L}$ applies.

To see how this definition applies, let's consider some examples. Let $\mathcal{L}$ is a list of relations that contains (exactly) a two-place relation R and a one place relation Q, then

- $(\forall x)(\forall y)(x = y)$ is not content-restricted to $\mathcal{L}$.
- $(\exists x)(Q(x) \wedge K(x))$ is not content-restricted to $\mathcal{L}$.
- $(\forall x)[x \in \text{Ext}(R) \rightarrow (\forall y)(y \in \text{Ext}(R) \rightarrow [R(x, y) \rightarrow Q(y)]])$[3] (which is first order logically equivalent to $(\forall x)(\forall y)[R(x, y) \rightarrow Q(y)])$ is content-restricted to $\mathcal{L}$.
- $\Diamond_R[(\forall x)(R(x, x) \wedge (\exists y)S(x, y))]$ is content restricted to $\mathcal{L}$

Also note the following consequences of the definition above:

- If $\mathcal{L}$ is a sublist of $\mathcal{L}'$, then all formulae ϕ which are content restricted

[3] i.e. $(\forall x)[(\exists k)(R(x, k) \vee R(k, x)) \rightarrow (\forall y)[(\exists k')(R(y, k') \vee R(k', y)) \rightarrow (R(x, y) \rightarrow Q(x))]]$

to $\mathcal{L}$ are also content restricted to $\mathcal{L}'$.

- A sentence is content restricted to the empty list $\mathcal{E}$ iff it is a truth functional combination of unsubscripted $\Box$ sentences, $\Diamond$ sentences and $\bot$.

As you may have noticed, explicitly content-restricted sentences are generally long and unwieldy. This can be annoying when writing up proofs whose inference steps can only (strictly speaking) be applied to sentences which are content-restricted to some list $\mathcal{L}$. To avoid this annoyance, I make the following definition.

Definition 11.3.2. A formula ϕ is **implicitly content-restricted** to $\mathcal{L}$ if there is a sentence ψ explicitly content restricted to $\mathcal{L}$ and $\phi \leftrightarrow \psi$ can be derived (using no assumptions) using only the first order inference rules, i.e., the principles already noted above.

I will then frequently use the shorthand of applying rules which (strictly speaking) can only be applied to content-restricted sentences to implicitly content restricted sentences – taking the work of using first order logic to deduce the explicitly content-restricted form of a sentence before applying the relevant rule (and then transforming it back after applying the rule) for granted.

I will say things like $\text{dom}(f) \subset Ext(R_1, ..R_n)$ to further abbreviate the claim that $(\forall x)\,(\text{dom}(f)(x) \rightarrow \text{Ext}(R_1, ...R_n)(x))$.

11.4 Content Restriction for Potentialist Paraphrases

With this definition of content restriction in hand, we see that our definitions of potentialist translations for set theoretic sentences and formulas are often content restricted in useful ways.

Lemma 11.4.1. *If $\phi, \theta_1, \ldots, \theta_n$ are formula in the language of set theory then*

1. *$t_n(\phi)$ is always content-restricted to $V_n, R_n, \mathbb{N}, \rho_n$*
2. *If ϕ is a sentence, then $t(\phi)$ is content restricted to the empty list.*
3. *For all i, j if $\vec{\mathscr{V}}(V_i), t_i(\theta_1), \ldots, t_i(\theta_n) \vdash_\Diamond t_i(\phi)$ then $\vec{\mathscr{V}}(V_j), t_j(\theta_1), \ldots, t_j(\theta_n) \vdash_\Diamond t_j(\phi)$*

Proof. Claims 1 and 2 follow immediately from the definition of content restriction and our potentialist paraphrases (repeated below). Claim 3 follows by a tedious, but simple, induction on proof length, where we transform the t_i version of a proof to the t_j version by replacing every instance of a relation in V_{i+k}, ρ_{i+k} with the corresponding relation V_{j+k}, ρ_{j+k} and noting that the result is still a proof. □

Chapter 12

The Formal System I: Basic Rules

12.1 Rules Inherited from First Order Logic

I will define the consequence relation $\vdash$ by listing closure conditions in this chapter and the next[1] Let $\Gamma, \Gamma_1, \Gamma_2$ be finite sets of formulas, and $\Gamma \vdash \theta$ express the claim that one can prove θ given the assumptions in Γ.

My closure conditions begin, boringly, with the following principles corresponding to standard inference rules for first order logic,(which I take from

[1]As usual, I will say that all variables that occur in an atomic formula are free. If a variable occurs free (or bound) in θ or in ψ, then that same occurrence is free (or bound) in $\neg\theta$, $(\theta \wedge \psi)$, $(\theta \vee \psi)$, and $(\theta \rightarrow \psi)$ and $\Diamond\theta$ and $\Box\theta$. That is, the (unary and binary) connectives do not change the status of variables that occur in them. All occurrences of the variable v in θ are bound in $(\forall v)\theta$ and $(\exists v)\theta$. Any free occurrences of v in θ are bound by the initial quantifier. All other variables that occur in θ are free or bound in $(\forall v)\theta$ and $(\exists v)\theta$, as they are in θ.

the Stanford Encyclopedia article on classical logic[96]).

(As) If ϕ is a member of Γ, then $\Gamma \vdash \phi$.

($\wedge$I) If $\Gamma_1 \vdash \theta$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash (\theta \wedge \psi)$.

($\wedge$E) If $\Gamma \vdash (\theta \wedge \psi)$ then $\Gamma \vdash \theta$; and if $\Gamma \vdash (\theta \wedge \psi)$ then $\Gamma \vdash \psi$.

($\vee$I) If $\Gamma \vdash \theta$ then $\Gamma_1 \vdash \theta \vee \psi$; if $\Gamma \vdash \psi$ then $\Gamma \vdash \theta \vee \psi$.

($\vee$E) If $\Gamma_1 \vdash (\theta \vee \psi)$, $\Gamma_2, \theta \vdash \phi$ and $\Gamma_3, \psi \vdash \phi$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \phi$.

($\rightarrow$I) If $\Gamma, \theta \vdash \psi$, then $\Gamma \vdash (\theta \rightarrow \psi)$.

($\rightarrow$E) If $\Gamma_1 \vdash (\theta \rightarrow \psi)$ and $\Gamma_2 \vdash \theta$, then $\Gamma_1, \Gamma_2 \vdash \psi$.

($\neg$I) If $\Gamma_1, \theta \vdash \psi$ and $\Gamma_2, \theta \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\theta$.

(DNE) If $\Gamma \vdash \neg\neg\theta$ then $\Gamma \vdash \theta$.

($\forall$E) If $\Gamma \vdash \forall v\theta$, then $\Gamma \vdash \theta(v|v')$, provided that v' is free for v in θ.[2]

($\forall$I) If $\Gamma \vdash \theta$ and the variable v does not occur free in any member of Γ, then $\Gamma \vdash \forall v\theta$.

(=I) $\Gamma \vdash v = v$, where v is any variable.

(=E) If $\Gamma_1 \vdash v_1 = v_2$ and $\Gamma_2 \vdash \theta$, then $\Gamma_1, \Gamma_2 \vdash \theta'$, where θ' is obtained from θ by replacing zero or more occurrences of v_1 with v_2, provided that no bound variables are replaced, and all substituted occurrences of v_2 are free.

[2] That is, if substituting v with v' does not lead to any variable which was antecedently free becoming bound. Here $\theta(v|v')$ stands for the result of substituting *all* free instances of v in θ with instances of v'.

($\perp$ I) If $\Gamma \vdash \psi \wedge \neg\psi$ then $\Gamma \vdash \perp$.

($\perp$ E) If $\Gamma, \theta \vdash \perp$ then $\Gamma \vdash \neg\theta$.

For convenience, I will also include the following inference rules for $\exists$ but, formally, I willl regard $\exists$ as an abbreviation for $\neg\forall\neg$.

($\exists$I) If $\Gamma \vdash \theta$, then $\Gamma \vdash \exists v\theta'$, where θ' is obtained from θ by substituting the variable v' for zero or more occurrences of a variable v, provided that (1) all of the replaced occurrences of v are free in θ, and (2) all of the substituted occurrences of v' are free in θ.

($\exists$E) If $\Gamma_1 \vdash \exists v\theta$ and $\Gamma_2, \theta \vdash \phi$, then $\Gamma_1, \Gamma_2 \vdash \phi$, provided that v does not occur free in ϕ, nor in any member of Γ_2.

I can now introduce the core axioms[3] which govern reasoning with $\Box$ and $\Diamond$ in my formal system.

Note that for the remainder of this book I adopt the convention that when used as meta-variables the capital greek letters $\Phi, \Psi, \Theta, \Xi, \Upsilon$ are restricted to sentences while the lower case greek letters $\phi, \psi, \theta, \xi, \upsilon$ may be formulas of sentences. I will also I follow the convention, standard in philosophical presentations of modal logic of calling these rules axioms, and presenting them in this form, even when they take the form of inference rules, i.e., closure conditions for $\vdash$ rather than sentences that can be inferred at any point.

[3]Here we follow the convention [37] of calling rules formulated in terms of formula meta-variables subject to syntactic constraints axioms rather than axiom schemas.

12.2 $\Diamond$ Introduction and Elimination

Axiom 12.2.1 ($\Diamond$ Introduction)**.**

$$\Theta \to \Diamond_{\mathcal{L}}\Theta$$

This rule captures the idea that what is actual must also be logically possible or, indeed, logically possible holding some facts about the actual world fixed.

This rule corresponds to rule T (sometimes written equivalently as $\Box A \to A$) in the familiar modal system S5 [37].

Examples:

- "There are two cats" $\Rightarrow$ "It is logically possible, given what cats there are, that there are two cats".
- "There are two cats" $\Rightarrow$ "It is logically possible, given what dogs there are, that there are two cats".

Axiom 12.2.2 ($\Diamond$ Elimination)**.** If Θ is content-restricted to $\mathcal{L}$ then

$$\Diamond_L\Theta \to \Theta$$

This rule expresses the idea that when θ is content-restricted to $\mathcal{L}$, the truth value of θ is totally determined by the facts about $\mathcal{L}$.

For instance:

- "It is logically possible, given what cats there are, that there are two cats" $\Rightarrow$ "There are two cats"
- BUT NOT: "It is logically possible, given what dogs there are, that there are two cats" $\Rightarrow$ "There are two cats"

Note that the second inference is not permitted by my rule because θ ("there are two cats") is not content-restricted to the list $\{dog(\cdot)\}$

The next basic axiom (schema) to expresses the intuition that if Θ is content restricted to $\mathcal{L}$ then holding fixed relations not in $\mathcal{L}$ doesn't affect the logical possibility of Θ.

12.3 $\Diamond$ Ignoring

Axiom 12.3.1 ($\Diamond$ Ignoring)**.** Suppose Θ is content-restricted to $\mathcal{L} = R_1, \ldots R_n$ and $S_1 \ldots S_m$ are relations not among $R_1, \ldots R_n$. Then $\Diamond_{\mathcal{L}'}\Theta \rightarrow \Diamond_{\mathcal{L}', S_1 \ldots S_m}\Theta$.

Remember that when a formula is content-restricted to $\mathcal{L}$, its truth depends only on facts about $\mathcal{L}$. This principle reflects this intuition by allowing one to ignore other facts.

We will see that the converse inference, from $\Diamond_{\mathcal{L}, S_1 \ldots S_m}\Theta$ to $\Diamond_{\mathcal{L}}\Theta$ is also provable from the basic axioms and inference rules in this chapter (in 13.1.1)

.

Examples:

- It is possible, given what cats there are, that there every cat admires a distinct dog → It is possible, given what cats and dolphins there are, that every cat admires a different dog.

- But NOT: It is possible, given what cats there are, that there are exactly 3 objects → It is possible, given what cats and dolphins there are, that there are exactly 3 objects.

  This inference is not permitted because the claim that there are exactly 3 objects is not content restricted to any list of relations including cats() but not dolphin().

- And NOT: It is possible, given what cats there are, that every cat admires a distinct dog → It is possible, given what cats and dogs there are, that every cat admires a distinct dog.

  Here θ is content restricted to {cat, dog, admires}, but for this inference to be permitted θ would have to be content restricted to a list that didn't include the relation dog since changing facts about what the relation dog() applies to could change the truth-value of θ.

12.4 Simple Comprehension

Axiom 12.4.1 (Simple Comprehension)**.** Suppose $R \notin \mathcal{L}$ and R isn't used by ϕ nor appears in Ψ then

$$\Psi \to \Diamond_{\mathcal{L}} (\Psi \wedge (\forall \vec{z}) [R(\vec{z}) \leftrightarrow \phi(\vec{z})])$$

This axiom schema captures the idea that any way a formula applies to objects is a logically possible way for a relation to apply to those objects. The inclusion of Ψ under the $\Diamond_{\mathcal{L}}$ reflects the intuition that the relation could apply in the way ϕ does without altering how any other relations apply or changing which objects exist[4].

12.5 Relabling

Axiom 12.5.1 (Relabeling)**.** If $R_1 \ldots R_n$ are distinct relations that don't occur in $\mathcal{L}$, and $R'_1 \ldots R'_n$ are distinct relations not equal to any R_i with the same arities[5] that don't occur in $\mathcal{L}$ or Θ , then

$$\Diamond_{\mathcal{L}}\Theta \leftrightarrow \Diamond_{\mathcal{L}}\Theta[R_1/R'_1 \ldots R_n/R'_n]$$

.

Here $\Theta[R_1/R'_1 \ldots R_n/R'_n]$ denotes the simultaneous substitution of R'_1 for R_1, R'_2 for R_2 and so on. Note that when we give a list of relations $R_1 \ldots R_n$ we usually assume they are distinct relations (no R_i is actually the same relation as R_j, $j \neq i$) but we state it explicitly here for clarity.

This axiom schema expresses the idea that when evaluating claims about logical possibility only the arity of a relation matters. Thus, replacing some $R \notin \mathcal{L}$ with an unused relation $R' \notin \mathcal{L}$ of the same arity cannot change the

[4]Technically, without changing the structure of the objects which are taken to exist at that 'world'.

[5]That is the arity of R_i and R'_i are the same.

truthvalue of $\Diamond_{\mathcal{L}}\theta$.

Example: By substituting sleeps with chews we see "It is logically possible, given the facts about dogs and blankets, that every dog sleeps on a different blanket" $\Leftrightarrow$ "It is logically possible, given the facts about dogs and blankets, that every dog chews on a different blanket."

> Example: "If there is something which everyone loves, it is logically possible (given the facts about love) that there is something which everyone loves *and* happy() applies to exactly those individuals which love themselves."

Note that unlike more familiar second order comprehension schemas, simple comprehension doesn't let infer that there is already a relation defined by the formula, only that it's possible. However, when it comes to proving results about what is possible, this difference imposes no further difficulty since, as we will see, we can collapse $\Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}$ into $\Diamond_{\mathcal{L}}$.

Note that as $\Box_{\mathcal{L}}$ is just $\neg\Diamond_{\mathcal{L}}\neg$ it is trivial to show that relabeling applies to $\Box$ claims as well as $\Diamond$ claims. We formalize this observation in the following lemma.

Lemma 12.5.2 (Box Relabeling)**.** *If $R_1 \ldots R_n$ are distinctrelations that don't occur in $\mathcal{L}$, and $R'_1 \ldots R'_n$ are distinct relations not equal to any R_i with the same arities as $R_1 \ldots R_n$ that don't occur in $\mathcal{L}$ and aren't mentioned in θ, then $\Gamma \vdash \Box_{\mathcal{L}}\theta \leftrightarrow \Box_{\mathcal{L}}\theta[R_1/R'_1 \ldots R_n/R'_n]$.*

12.6 Importing

Axiom 12.6.1 (Importing)**.** If Θ is content restricted to $\mathcal{L}$ then

$$[\Theta \wedge \Diamond_{\mathcal{L}}\Phi] \rightarrow \Diamond_{\mathcal{L}}(\Phi \wedge \Theta)$$

This rule captures the idea that any true sentence Θ which only talks about how some relations $\mathcal{L}$ apply, must remain true in any logically possible context that holds the $\mathcal{L}$ facts fixed.

12.7 Logical Closure

Inference Rule 12.7.1 (Logical Closure)**.** If $\Theta \vdash \Phi$ then

$$\Diamond_{\mathcal{L}}\Theta \rightarrow \Diamond_{\mathcal{L}}\Phi$$

Note that, despite being called axioms, the principles in this chapter and the next need not be listed in the premises of an argument, e.g., if Θ is content restricted to $\mathcal{L}$ then, by $\Diamond$ Elimination,

$$\emptyset \vdash \Diamond_{\mathcal{L}}\Theta \rightarrow \Theta$$

This rule captures the idea that logical inference is universally valid. Thus, if we can deduce Φ from Θ then a 'scenario' in which Θ is true must also be one in which Φ is true. Note that when $\mathcal{L}$ is empty, this rule performs the work of both the necessitation and distribution rules commonly used of in systems like S5.

Chapter 13

Notation and Some Example Arguments

Before we present any more results in the system it is worth remarking on the status of relations that appear in a lemma. For instance, consider the trivial lemma whose content is $(\exists x)R(x) \to (\exists x)R(x)$. We don't regard the lemma as merely a proof of the fact that for some particular relation, e.g., red, if there is some red thing then there is some red thing. Rather, we regard the lemma as standing in for the fact that this result is provable for *any* one place relation or, alternately, as proving that the claim in the lemma is logically necessary[1]. Thus, we allow deducing the fact that $(\exists x)G(x) \to (\exists x)G(x)$ from the trivial lemma asserting that $(\exists x)R(x) \to (\exists x)R(x)$. Note that this resembles the situation in first-order logic where we prove that substitution

[1]As we will see shortly if we can prove $(\exists x)R(x) \to (\exists x)R(x)$ without any premises we can infer $\Box(\exists x)R(x) \to (\exists x)R(x)$ and then (as we will also see below) substitute the relations under the $\Box$ and then eliminate it.

of bound variables preserves truth-value and then don't pay much attention to the particular bound variables used to express results.

13.1 $\Diamond$ Reducing and $\Box$ Expansion

Using first order logic (FOL) and the basic principles above we can prove various useful lemmas.

The Reducing Lemma (together with $\rightarrow$E) vindicates intuitive reasoning along the following lines. Suppose it's logically possible, given the facts about friendship and enmity in the actual world, that something has a frenemy (i.e., there are items x and y such that x is the friend of y and x is the enemy of y). Then it's logically possible given (just) the facts about friendship in the actual world that something has a frenemy.

Lemma 13.1.1 (Reducing)**.** *If* $\mathcal{L} \supseteq \mathcal{L}'$ *then*

$$\Diamond_{\mathcal{L}}\Theta \rightarrow \Diamond_{\mathcal{L}'}\Theta$$

Proof. First note that if $\mathcal{L} \supseteq \mathcal{L}'$ then any sentence of the form $\Diamond_{\mathcal{L}'}\Theta$ is content restricted to $\mathcal{L}$.

Assume that $\Diamond_{\mathcal{L}}\Theta$. We have $\Theta \vdash \Diamond_{\mathcal{L}'}\Theta$, by $\Diamond$ Introduction. So, by Logical Closure, we have $\Diamond_{\mathcal{L}}(\Diamond_{\mathcal{L}'}\Theta)$. Then by $\Diamond$ Elimination we can conclude that $\Diamond_{\mathcal{L}'}\Theta$ (since $\Diamond_{\mathcal{L}'}\Theta$ is content restricted to $\mathcal{L}$). Thus we have $\Diamond_{\mathcal{L}}\Theta \rightarrow \Diamond_{\mathcal{L}'}\Theta$ $\Box$

We note that this immediately entails a corresponding expansion property for sentences under the $\Box$.

Lemma 13.1.2 (Box Expanding)**.** *Suppose* $\mathcal{L}' \supset \mathcal{L}$ *then*

$$\Box_{\mathcal{L}}\Phi \rightarrow \Box_{\mathcal{L}'}\Phi$$

Proof. Suppose the claim fails. Then we have $\Box_{\mathcal{L}}\Phi$ and $\Diamond_{\mathcal{L}'}\neg\Phi$. By Reducing we can infer $\Diamond_{\mathcal{L}}\neg\Phi$ which contradicts the assumption above. □

13.2 Inner Diamond

We will can also prove the Inner Diamond lemma, which will help us capture natural reasoning about conditional logical possibility in a more natural manner. Specifically, while the Importing and Logical Closure axioms capture the intuition that we can deploy our normal tools of reasoning to infer what further facts must be true in some particular logically possible context using them directly would force us to carry unwieldy long conjunctions of all facts we've derived are logically possible through our proofs. The Inner Diamond lemma justifies our use of more natural mathematical reasoning.

The intuition behind the Inner Diamond lemma is that reasoning like the following is valid.

> Suppose we know the following. There are at least three cats. And it's logically possible, given what cats there are, that every cat is sleeping on a distinct blanket. What else must be true in this logically possible scenario? We can 'import' the fact that there are at least three cats (since any scenario which preserves

> the structural facts about how cathood applies must preserve this fact). So, by first order logic, this possible scenario must be one in which there are at least four blankets. Thus it is logically possible, given the facts about what cats there are, that there are at least three blankets.

Proposition 13.2.1 (Inner Diamond). *If* $\Gamma_1 \vdash \Diamond_{\mathcal{L}}\Theta$ *and* $\Gamma_2, \Theta \vdash \Phi$*, where every element of* Γ_2 *is a sentence content-restricted to* $\mathcal{L}$*, then* $\Gamma_1, \Gamma_2 \vdash \Diamond_{\mathcal{L}}\Phi$*.*

Proof. Consider a scenario where the antecedent of the lemma is true. Assume that Γ_1, Γ_2. Then we have $\Diamond_L\Theta$ by the first assumption. By successive applications Importing to each of the sentences $\Gamma_2^1, \ldots, \Gamma_2^n$ in Γ_2 , we have $\Diamond_L\Theta \wedge \Gamma_2^1 \wedge \ldots \wedge \Gamma_2^n$. Now by Logical Closure and the fact that $\Gamma_2, \Theta \vdash \Phi$ we can get $\Diamond_L\phi$. Thus $\Gamma_1, \Gamma_2 \vdash \Phi$ as desired. □

We note that this lemma supports the following kind of reasoning (as illustrated in the above example). We derive some sentence of the form $\Diamond_{\mathcal{L}}\Theta$ from the assumptions Γ_1. For instance in the example above Θ would be the claim that 'every cat slept on a distinct blanket' and $\mathcal{L}$ would just be the predicate cat. We then wish to reason about the 'world' whose possibility is guaranteed by the fact that $\Diamond_{\mathcal{L}}\Theta$, e.g., the world which holds fixed (the structure of) the application of cat and at which every cat slept on a distinct blanket. In that world Θ (every cat slept on a distinct blanket) is true as is, intuitively, every fact content restricted to {cat} true in the actual world. For instance, in the example above the fact that there are at least three cats is also true in that world (we refer to the act of taking a sentence content restricted to $\mathcal{L}$ and concluding it holds at the world whose logical possibility is asserted by $\Diamond_{\mathcal{L}}\Theta$ as importing). We then use proof rules to derive some desired conclusion Φ from Θ and the set of 'imported' sentences Γ_2. For instance in the above example Φ is the sentence asserting there are at least 3 blankets. Intuitively, Φ must also be true in the logically possible world under consideration and thus $\Diamond_{\mathcal{L}}\Phi$ must be actually true. In the example above Γ_2 would just contain the sentence asserting that there are at least three cats. Since $\Theta, \Gamma_2 \vdash \Phi$ and all sentences in Γ_2 are content restricted to $\{\text{cat}\}$ this intuition is born out rigorously since the above lemma establishes that $\Gamma_1, \Gamma_2 \vdash \Diamond_{\mathcal{L}}\Phi$.

13.3 Natural Deduction With Inner Diamond Arguments

Since the process of entering $\Diamond_{\mathcal{L}}$ contexts, i.e., using Inner Diamond to reason about what else must be true in a particular logically possible scenario, is unfamiliar and can be a bit tricky I will informally introduce a natural deduction system for the notion of proof defined chapter 12 and chapter 14 together with some notational conventions which make it easier to keep track of arguments like the one above (especially in contexts where one must make multiple inner diamond arguments within one another).

This system is loosely based around that used in [39] and I follow his system in citing the line numbers justifying each inference rule to the left of the name of the inference rule, while indicating the assumptions a line depends on by placing those line numbers in brackets (line numbers not in brackets are the lines cited as immediate justification for the current inference). So, for example, we write down $\Phi \quad 5,6 \quad X \quad [2,4,5]$ on line 7 of a proof when rule X allows us to conclude Φ from lines $5,6$ and the cumulative set of assumptions from which we've established Φ are the sentences on lines $2,4$ and 5. Note that this system satisfies the principle that if ψ appears on some line of the proof and Γ is the set of sentences appearing on the lines listed in brackets next to ψ then $\Gamma \vdash \psi$. However, my system differs from his in two primary ways.

First, I will allow any purely first order deduction to be compressed into a single FOL rule. However, I will still sometimes explicitly make use of $\rightarrow I$ to infer $\phi \rightarrow \psi$ in cases where ψ can only be inferred from ϕ via modal reasoning. I will also use Ass. to indicate that a new assumption is being made.

Second, all modal axioms and axiom schema proposed in chapters 12 and 14 are taken to be logical truths. So any instance of these axiom schemata can be written down with no associated citations or assumptions. And, to save time, any instance of an axiom schema with the form $\phi \rightarrow \psi$ may instead be regarded as an inference rule allowing us to infer ψ from ϕ (citing the line containing ϕ as a justification). For example, this is an acceptable deduction of $\Diamond_P [(\exists x)P(x)] \rightarrow P(x)$.

1	$\Diamond_P [(\exists x)P(x)] \rightarrow P(x)$	$\Diamond$ E

This is also an acceptable deduction of the same fact.

1	$\Diamond_P [(\exists x)P(x)]$	Ass [1]
2	$P(x)$	1 $\Diamond$ E [1]
3	$\Diamond_P [(\exists x)P(x)] \rightarrow P(x)$	1,2 $\rightarrow$ I

Third, and most distinctively, I will introduce a special context called a $\Diamond$ context (nestable to arbitrary depth) corresponding to reasoning via Inner Diamond, i.e., reasoning about what else must be true within some scenario which is known to be (conditionally) logically possible. I will graphically indicate what sentences are being asserted or assumed within this context by indentation and a sideways T labeled with a $\Diamond$ to indicate this context.

So, for example, we can represent the following extremely short inner diamond argument

> Given what cats and hunters there are, it's logically possible that something is both a cat and a hunter. Any possible situation in which something is a cat and a hunter must be a situation in which something is either a cat or a hunter. Therefore, given what cats and hunters there are, its logically possible that something is either a cat or a hunter.

with a proof that looks like this

1	$\Diamond_{cat}(\exists x)(cat(x) \wedge hunter(x))$	[1]
2	$\Diamond$ $\mid$ $(\exists x)(cat(x) \wedge hunter(x))$ $[cat, hunter]$	(1) In$\Diamond$ I, [2*]
3	$\mid$ $(\exists x)(cat(x) \vee hunter(x))$	(2) FOL [2*]
4	$\Diamond_{cat}(\exists x)(cat(x) \vee hunter(x))$	(1, 2-3) In$\Diamond$ E [1]

The vertical line going from 2-3 above indicates those lines occur inside a special context. I call this a $\Diamond$ context to indicate that these lines contain reasoning about *what must be true within* a logically possible scenario in which $(\exists x)(cat(x) \wedge hunter(x))$, while all the structural facts about, how cathood applies, are preserved.

What are the rules for writing things down in this context? Recall that the Inner Diamond lemma says that if we have one conditional possibility claim $\Diamond_{\mathcal{L}}(\Theta)$, and some facts Γ_2 which are content restricted to the relations being held fixed, then if we can show that any such possible scenario where Θ must also be one where Φ (by showing $\Theta, \Gamma_2 \vdash \Phi$), we can infer the corresponding conditional logical possibility clam for Φ.

The key idea will be to use indentation and the Fitch style sidewise 'T' to graphically distinguish a main line of argument which goes from Γ_2 (where the sentences Γ_2 are content restricted to $\mathcal{L}$) and $\Diamond_{\mathcal{L}}\Theta$ and then $\Diamond_{\mathcal{L}}\Phi$,from a supporting subproof which shows that $\Gamma_2, \Theta \vdash \Phi$ and thereby justifies the latter inference.

In the latter subproof (which I indent and mark off as a separate context) we are, in essence, attempting to milk new consequences from our knowledge that $\Diamond_{\mathcal{L}}(\Theta)$, by thinking about what *else* must be true in a possible ($\Diamond_{\mathcal{L}}$) stuation where Θ. Thus I will call beginning such a subproof 'entering the $\Diamond_{\mathcal{L}}$ context' (associated with some previous claim that $\Diamond_{\mathcal{L}}(\Theta)$)', and thereby beginning an Inner Diamond argument.

One will only be permitted to import some claims Γ_2 from the main line of argument (thereby assumeing they continue to hold in the current context) if these claims are content restricted to the relevant list of relation $\mathcal{L}$. And we will only be allowed to 'close the inner diamond context' by dropping back one level of indentation and writing down $\Diamond_{\mathcal{L}}(\Phi)$ if we have proved that Φ must hold in this context, by showing that it follows from the initial assumption that Θ and some facts Γ_2 which we have imported via general logical laws.

Reasoning inside a $\Diamond$ context proceeds just as it does normally, with the exception that each line in the context must either be our initial assumption that Θ (where $\Diamond_{\mathcal{L}}(\Theta)$ is the sentence that opened the diamond context), an instance of 'importing' (where the sentence must be imported from the parent context) or be deducible from previous lines within this exact $\Diamond$ context.

While the operation of In$\Diamond$I is pretty straightforward we call attention to one detail. Note that besides line 2 we wrote [2*] rather than [1] as one might expect. We do this to maintain the property that if Ψ is written on a line it is deducible from the lines written in brackets next to it[2].

One can leave the $\Diamond_{\mathcal{L}}$ context above by going from knowledge that ϕ holds within this context to the conclusion that $\Diamond_{\mathcal{L}}\phi$ holds outside it. We indicate this inference pattern via the rule In$\Diamond$E. This is the only way to introduce a sentence into the current context based on activity in a child context[3].

Example of Inner $\Diamond$ with importing:

We can also capture the reasoning in the slightly more complicated argument below, where we use knowledge of sutiably content restricted claims about the acutal world to draw consequences from a modal claim.

It's logically possible, given what cats there are, that each slept

[2]To this end we treat the initial line in each $\Diamond$ context and every line introduced via importing (see below) as if they were assumptions inside that context. However, we mark these assumptions with an asterisk since they are justified assumptions (it's safe to assume they are true in the $\Diamond$ context) and must be replaced with the line numbers from the parent context when we leave the $\Diamond$ context.

In accordance with this idea, a sentence ρ can be written down inside the "$\Diamond_{\mathcal{L}}$ context" governed by the claim that $\Diamond_{\mathcal{L}}\Theta$, iff

- $\rho = \Theta$
- $\rho = \Psi$ for some Γ which is content-restricted to $\mathcal{L}$ and occurs on an earlier line in the proof which is in the same context as the $\Diamond_{\mathcal{L}}\Theta$ statement used to introduce this inner diamond context. (I will, as usual, sometimes elide the steps needed to transform implicitly content restricted sentences into first order logically equivalent explicitly content restricted sentences.)
- ρ follows from previous lines within this $\Diamond$ context by one of the axioms or inference rules for reasoning about logical possibility presented in this book.

[3]Note that In$\Diamond$E may not be applied to any line with uncancelled (unstarred) assumptions introduced in the context being closed. Moreover, In$\Diamond$E must take each starred line number j^* on the line on which Φ appears (here that's line 3 and Φ is$(\exists x)(cat(x) \vee hunter(x))$) and replace it with the assumptions of the line (in the current context) used to justify line j. [take this for inst and put duplicate of it below in example with importing] For instance, in the current case the only (starred) assumption for line 3 is line 2. Looking at line 2 we see that it is justified by reference to line 1 (which is in the current context). So we copy the line numbers in brackets on line 1 into the brackets on line 4 (in this case that's just 1).

on a distinct blanket. There are at least three cats. Therefore, it's logically possible, given what cats there are that there are at least three blankets.

1	$\Diamond_{cat}$Each cat slept on a distinct blanket	Ass. [1]
2	There are at least three cats	Ass. [2]
3	$\Diamond$ | Each cat slept on a distinct blanket	1 In$\Diamond$, [3*]
4	| There are more than three cats	2 Import [4*]
5	| There are at least three blankets	3,4 FOL [3*,4*]
6	$\Diamond_{cat}$There are at least three blankets	1, 2-5 In$\Diamond$E [1,2]

Note that the sentence on line 2 'There are at least three cats' is content restricted to $\{cats\}$ (assuming that this abbreviates an FOL statement in the usual Fregian fashion). This fact allows us to import it into our reasoning about what the possible scenario where each cat slept on a different blanket must be like on line 4.

Also note that on line 5 we have proved 'there are at least blankets' with only assumptions $[3^*, 4^*]$ (which are starred because they were introduced by inner diamond introduction or importing). Thus, we have shown that the conclusion that there are more than three blankets follows from the things we are entitled to assume about any logically possible scenario witnessing the truth of the sentence on line 1. So we can apply Inn$\Diamond$ Elimination to complete our inner diamond argument, and conclude that $\Diamond_{cat}$There are more than three blankets.

Finally, note that the assumption line numbers listed for the our conclusion are $[1, 2]$. For these are the assumptions needed for the claims about the actual world (namely the sentences on lines 1 and 2), which entitle us to assume that the possible scenario considered on lines 3-5 satisfy the assumptions on lines 3 and 4 which imply are more than three blankets.

See Appendix C for an explicit formal statement of this natural deduction system, and a demonstration that proofs in it obey the notion of provability above.

13.4 Box Inference Rules

Although the $\Box$ is not an official item in our symbolism, but merely an abbreviation for $\neg\Diamond\neg$, it is often helpful to reason in terms of it. Earlier we proved a couple of rules regarding $\Box$ inferences and here we present several more.

First, I present an introduction rule for $\Box$.

Lemma 13.4.1 ($\Box$ I)**.** *If* $\Gamma \vdash \Theta$ *and every* $\gamma \in \Gamma$ *is a sentence content-restricted to* $\mathcal{L}$ *then* $\Box_{\mathcal{L}}\theta$*.*

Proof. Suppose, for a contradiction, Γ, Θ are as above but the lemma fails, i.e., $\Diamond_{\mathcal{L}}\neg\Theta$. By Inner Diamond with $\Theta_1 = \{\Diamond_{\mathcal{L}}\neg\Theta\}$ and $\Theta_2 = \Theta$ we can infer $\Diamond_{\mathcal{L}}\bot$ as $\Theta, \neg\Theta \vdash \Theta \wedge \neg\Theta$. Hence, by $\Diamond$ Elimination we can export the contradiction. Hence, $\Box_{\mathcal{L}}\Theta$ as desired.

$\square$

Now I give the corresponding elimination rule.

Lemma 13.4.2 ($\Box$ Elimination)**.**

$$\Box_{\mathcal{L}}\Theta \rightarrow \Theta$$

Note that in I prove a stronger version of this result in appendix D that allows arbitrary substitution of relations when eliminating the box and, for ease of reading, I will also refer to that result as $\Box$ Elimination.

Proof. Assume the claim fails. We can derive contradiction immediately by applying $\Diamond$ Introduction to $\neg\Theta$ to derive $\Diamond_{\mathcal{L}}\neg\Theta$ which is $\neg\Box_{\mathcal{L}}$. We can write this in terms of the natural deduction system presented above as follows.

1	$\Box_{\mathcal{L}}\Theta$	[Γ]
2	$\neg\Diamond_{\mathcal{L}}\neg\Theta$	[Γ]
3	$\neg\Theta$	Assump. [3]
4	$\Diamond_{\mathcal{L}}\neg\Theta$	4 $\Diamond$ I [3]
5	$\bot$	2, 4 $\bot$ I [3,Γ]
6	$\neg\neg\theta$	3-5 ¬I [Γ]
7	θ	6 ¬E [Γ]

□

13.5 Lemmas about Well-Orderings

To give a more visceral sense of how proofs using this logical system work I will now prove two lemmas which mirror results in set theory (which can be found in elementary texts like [50]). In each case, I will make an argument verbally, and then follow it up with an argument using the formal notation (making explicit when we enter and leave Inner Diamond arguments).

In later chapters and the appendixes I will present proofs in a more informal style. However, I hope the proofs in this chapter will help the reader understand how these informal proofs can be expanded into a formal argument.

13.5.1 Lemma A

Jech's version of the first lemma I am going to prove says the following:

> "If $(W, <)$ is a well-ordered set and $f : W \to W$ is an increasing function, then $x < f(x)$ for each $x \in W$."[50]

We can write a version of Jech's Lemma follows (see appendix A for the definition of a well-order):

Lemma 13.5.1. *If f is an embedding of the well-order $W, <$ into itself then*

$$(\forall x, y : W(x) \wedge W(y))(x < f(x))$$

where we define

Definition 13.5.2 (Definition of Embedding)**.** A two place relation f **is an embedding** of $W, <$ into $W', <'$ iff

- f is a function (remember we define what it takes for a relation to qualify as a function in section 8.1)
- $(\forall x)[W(x) \to (\exists y)(W'(y) \wedge f(x)]$ i.e., f maps all of W into W'
- $(\forall x)(\forall y)(\forall x')(\forall y')[x < y \leftrightarrow f(x) < f(y)]$, i.e., f respects $<$.

Remember that we've defined function so that the function $f(x)$ is a convenient way of talking about the relation $f(x, y)$ satisfying $f(x, y) \wedge f(x', y) \to x = x'$.

As usual, I will sometimes abbreviate the claim that $x < y \vee \neg x = y$ as $x \leq y$.

Proof. To prove this, we will use essentially the same reasoning which Jech uses to prove his set theoretic version of this claim.

Assume that f is an embedding of $(W, <)$ into itself, as per the statement of the lemma.

And suppose, for contradiction, the lemma fails. As in Jech's proof, our aim will be to use the properties of well-orderings to derive the existence of a $<$ least counterexample, i.e., an x in W such that $\neg x < f(x) \wedge (\forall y : y < x)(y < f(y))$ and derive contradiction from this.

Applying Simple Comprehension to the formula below

$$\neg x < f(x)$$

tells us it would logical possible - while holding fixed the facts about how $W, <, f$ apply in the situation we are currently considering - for the predicate G apply to just such counterexamples. That is

$$\Diamond_{W,<,f}(\forall x)(G(x) \leftrightarrow \neg x < f(x))$$

Now we can enter this $\Diamond_{W,<,f}$ context, i.e., begin an Inner Diamond argument, where we reason about what *else* must be true in a possible scenario where (the facts about $W, <, f$ in our original scenario are held fixed but) we also have:

$$(\forall x)(G(x) \leftrightarrow \neg x < f(x)) \tag{13.1}$$

Now the premises of the lemma (that f is an embedding of $(W, <)$ into itself and $(W, <)$ a well order) and the assumption that the conclusion of the lemma fails are all implicitly context restricted to $\{W, <, f\}$ (seen by appropriately restricting all the quantifiers). So all of these statements must all remain true in this new context and can by imported into this context.

Thus we can infer that G is non-empty, from the assumption that the lemma fails, i.e., $(\exists x \mid W(x))(\neg x < f(x)$, together with the fact that $(\forall x)(G(x) \leftrightarrow \neg x < f(x))$.

We know that $W, <$ is a well ordering, and the least element condition from the definition of well ordering (A.0.2) says the following:

$$\Box_{W,<} [(\exists x \mid W(x))\, G(x) \rightarrow (\exists y \mid W(y) \wedge G(y))\, (\forall z \mid W(z) \wedge G(z))\, (y \leq z))]$$

So by $\Box$ Elimination we can infer the existence of a least counterexample y, i.e.,

$$(\exists y \mid W(y) \wedge G(y))\, (\forall z \mid W(z) \wedge G(z)))\, (y \leq z))$$

Now let $z = f(y) < y$. By our assumption that f is an embedding (and thus must respect $<$) it follows from $z < y$ that $f(z) < f(y) = z$. So by eq. (13.1) $G(z)$. Thus, z is an satisfyies $G(z)$ and is less than y. Contradiction $\perp$.

Exiting the above $\Diamond$ context (i.e., completing our Inner Diamond argument), we get:

$\Diamond_{W,<,f}\perp$

And from this $\perp$ follows by $\Diamond$ Elimination (remembering that $\perp$ is content restricted to the empty list). Hence, the desired conclusion follows by contradiction. $\square$

Intuitively speaking, the argument above shows that the if there were a counterexample to the lemma then it would be logically possible (indeed logically possible, while holding fixed the $W, <, f$ facts!) for the canonical contradiction $\perp$ to be true. But it's not logically possible for $\perp$ to be true. So there is no counterexample to the lemma.

Representing this proof in terms of our natural deduction system.

1	f is an embedding of the well-order $(W, <)$ into itself	[1]
2	$(\exists x \mid W(x))(f(x) < x)$	[2]
3	$\Diamond_{W,<,f} \forall z[G(z) \leftrightarrow f(z) < z]$	3 Simple Comprehension
4	$\Diamond$ | $\forall z[G(z) \leftrightarrow f(z) < z]$ $\{W, <, f\}$	3 In$\Diamond$ I [4*]
5	| f is an embedding of the well-order $W, <$ into itself	1 imp [5*]
6	| $(\exists x \mid W(x))(f(x) < x)$	2 imp [6*]
7	| $\Box_{W,<}[(\exists x \mid W(x) \wedge G(x)) \rightarrow (\exists y)[W(y) \wedge G(y) \wedge (\forall z)(W(z) \wedge G(z) \rightarrow y \leq z)])]$	5 FOL [5*]
8	| $(\exists x \mid W(x) \wedge G(x)) \rightarrow (\exists y)[W(y) \wedge G(y) \wedge (\forall z)(W(z) \wedge G(z) \rightarrow y \leq z)])$	7 $\Box$ E [5*]
9	| $\bot$	4, 6, 8 FOL [4*,5*,6*]
10	$\Diamond_{W,<,f} \bot$	3, 4-9 In$\Diamond$E [1,2]
11	$\bot$	10 $\Diamond$E [1,2]
12	$\neg(\exists x)(f(x) < x)$	11 (2) FOL [1]

13.5.2 Lemma B

Jech writes

"No well-ordered set is isomorphic to an initial segment of itself"[51]

We can define the claim that two structures $R_1, \ldots, R_n$ and $R'_1, \ldots, R'_n$ are isomorphic modally, as saying that it's possible for a relation f to map the objects in (the extension of) one structure to those in (the extension of) the other structure.

Definition 13.5.3 (Isomorphism)**.** A relation f is an isomorphism of $\langle R_1, \ldots R_m \rangle$ with the $\langle R'_1, \ldots R'_m \rangle$ (henceforth written $\langle R_1, \ldots R_m \rangle \cong_f \langle R'_1, \ldots R'_m \rangle$) if:

- f is a bijection of $\text{Ext}(R_1, \ldots R_m)$ with $\text{Ext}(R_1, \ldots R_m)$ (note that the domain of f may be larger than $\text{Ext}(R_1, \ldots R_m)$ so long as it behaves appropriately on $\text{Ext}(R_1, \ldots R_m)$).

- f respects the relations R_i, R'_i, i.e.,

$$(\forall x_1) \ldots (\forall x_m) \left[R_i(x_1 \ldots x_n)) \leftrightarrow R'_i(f(x_1) \ldots f(x_n)) \right]$$

  ,

So we can state the claim to be proved as follows.

Lemma 13.5.4. *If* $(W, <)$ *is a well-ordering and there is some* x *in* W *such that* W' *applies to just those* $z < x$ *in* W *then*

$$\neg \Diamond_{W,W',<} \langle W, > \rangle \cong_f \langle W', > \rangle$$

Proof. Let $W, W', <$ be as in the lemma and suppose for contradiction that $\Diamond_{W,W',<} \langle W, > \rangle \cong_f \langle W', > \rangle$. Using Inner Diamond we can enter this $\Diamond$ context. We can import the fact $W, <$ is a well order (because it is content-restricted to W, W' and $<$). By first order logic and unpacking definitions we can infer from the fact that f isomorphically maps $\langle W, > \rangle$ to $\langle W', > \rangle$

that f is an embedding of $W, <$ into $W', <$. And, by the assumptions about W' above, this implies that f is an embedding of $W, <$ into $W, <$.

Now, to get contradiction, note that by Lemma A (all instances of which are provable from empty premises, hence provable in any $\Diamond$ context) f does not map any object satisfying W strictly $<$-below itself. However, we know there is an object x satisfying W which is $>$ all objects satisfying W' and that $\langle W, > \rangle \cong_f \langle W', > \rangle$. It follows by first order logic that f maps this x to a some $y < x$. Thus we have derived contradiction/the false ($\perp$) from premises which would have to obtain in this (supposedly) logically possible scenario.

As before, we can conclude this inner $\Diamond_{W,W',<}$ argument and returning to our original context with the conclusion that $\Diamond_{W,W',<}\perp$. And from this $\perp$ follows by $\Diamond$ Elimination.

This completes our proof by contradiction that there can be no f isomorphicly mapping $(W, >)$ to a proper initial segment of itself.

□

We can use the natural deduction system to expose the modal reasoning within this argument, as follows.

1	$(W,<)$ is a well ordering and W' is a proper initial segment of W	Ass. [1]
2	$\Diamond_{W,W',<}\langle W,>\rangle \cong_f \langle W',>\rangle$	Ass. [2]
3	$\Diamond$ | $\langle W,<\rangle \cong_f \langle W',<\rangle \qquad \{W,W',>\}$	2 In$\Diamond$I [3*]
4	| $(W,<)$ is a well ordering and W' is a proper initial segment of W	1 import [4*]
5	| f is an embedding of the well-order $(W,<)$ into itself	3,4 FOL [3*,4*]
6	| $\neg(\exists x)(f(x) < x)$	6 Lemma A
7	| $\bot$	3,4,6 FOL[3*,4*]
8	$\Diamond_{W,W',<}(\bot)$	2,3-7 In$\Diamond$E [1,2]
9	$\bot$	8 $\Diamond$E [1,2]
10	$\neg\Diamond_{W,W',>}\langle W,>\rangle \cong_f \langle W',>\rangle$	9 (2) FOL [1]

13.6 Pasting and Collapsing

Finally, I will conclude this chapter with two lemmas involving of how more complex modal reasoning involving multiple $\Diamond$ contests. The first lemma tells us when two logically possible facts can be inferred to be jointly possible.

One cannot generally infer from $\Diamond_{\mathcal{L}}\Phi$ and $\Diamond_{\mathcal{L}}\Psi$ to $\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$. For consider the case where Φ says there are exactly 8 million things and Ψ says there are exactly 9 million things. However, the Pasting Lemma says that one *can* make this inference in the special situation when the sentences Φ and Ψ are content restricted so that they can only make claims about suitably disjoint aspects of relation (and how these relate the actual $\mathcal{L}$-structure, which both $\Diamond_{\mathcal{L}}\Phi$ and $\Diamond_{\mathcal{L}}\Psi$ preserve).

Lemma 13.6.1 (Pasting)**.** *Let $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{L}$ be pairwise disjoint sets of relations. If $\Diamond_{\mathcal{L}}\Phi$, where Φ is content restricted to $\mathcal{L},\mathcal{I}$ and $\Diamond_{\mathcal{L}}\Psi$, where Ψ is content-restricted to $\mathcal{L},\mathcal{J}$, then $\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$.*

Intuitively speaking, the facts about content restriction above ensure that attempting to make the sentences inside both possibility claims true at the same time cannot impose conflicting demands. For the only relations whose extensions are relevant to the truth of both sentences are map and adjacent to. And our assumptions say that it's possible to make each interior sentence true while fixing the actual application of these relations.

Proof. Let Φ be content restricted to $\mathcal{L},\mathcal{I}$ and Ψ to $\mathcal{L},\mathcal{J}$, as per the antecedent.

1	$\Diamond_{\mathcal{L}}\Phi$	Ass. [1]
2	$\Diamond_{\mathcal{L}}\Psi$	Ass. [2]
3	$\Diamond$ Φ $\{\mathcal{L}\}$	1, In$\Diamond$I [3*]
4	$\Diamond_{\mathcal{L}}\Psi$	2 import [4*]
5	$\Diamond_{\mathcal{L},\mathcal{I}}\Psi$	4 $\Diamond$ Ignoring [4*]
6	$\Diamond$ Ψ $\{\mathcal{L},\mathcal{I}\}$	5 In$\Diamond$I [6*]
7	Φ	3 import [7*]
8	$\Phi \wedge \Psi$	6,7 FOL [6*,7*]
9	$\Diamond_{\mathcal{L},\mathcal{I}}(\Phi \wedge \Psi)$	5,6-8 In$\Diamond$E [3*,4*]
10	$\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$	9 Reducing [3*,4*]
11	$\Diamond_{\mathcal{L}}(\Diamond_{\mathcal{L}}(\Phi \wedge \Psi))$	1,3-10 In$\Diamond$E [1,2]
12	$\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$	11 $\Diamond$E [1,2]

Informally, this deduction corresponds to the following reasoning:

Assume that $\Diamond_{\mathcal{L}}\Phi$ and $\Diamond_{\mathcal{L}}\Psi$. We can prove our claim by making two nested Inner Diamond arguments.

First enter the ($\Diamond_{\mathcal{L}}$) context associated with $\Diamond_{\mathcal{L}}\Phi$. That is, consider what else must be true in any such possible ($\Diamond_{\mathcal{L}}$) situation where Φ. In this situation $\Diamond_{\mathcal{L}}\Psi$ must remain true, for it is content restricted to $\mathcal{L}$, and we are considering a scenario which preserves the $\mathcal{L}$ facts. By $\Diamond$ Ignoring it follows that $\Diamond_{\mathcal{L},\mathcal{I}}\Psi$.

Now enter this second, interior, $\Diamond_{\mathcal{L},\mathcal{I}}$ context. That is, consider what must be true in a further possible scenario where Ψ is true while all facts about how relations $\mathcal{L},\mathcal{I}$ applied in the scenario we previously considered are preserved. Here we clearly have Ψ. But we can import the fact that Φ from the previous context, because it is content restricted to $\mathcal{L},\mathcal{I}$. So we can deduce $\Phi \wedge \Psi$.

Now, leaving this inner $\Diamond_{\mathcal{L},\mathcal{I}}$ context, we can conclude that $\Diamond_{\mathcal{L},\mathcal{I}}(\Phi \wedge \Psi)$. And we can infer that $\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$ by $\Diamond$ Ignoring (because $\mathcal{L}$ is clearly a sublist of $\mathcal{L},\mathcal{I}$).

So, leaving the larger $\Diamond_{\mathcal{L}}$ context we can conclude that $\Diamond_{\mathcal{L}}(\Diamond_{\mathcal{L}}(\Phi \wedge \Psi))$ holds in the situation we were originally considering.

Finally, because $\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$ is content restricted to $\mathcal{L}$, we can use $\Diamond$E to draw the desired conclusion $\Diamond_{\mathcal{L}}(\Phi \wedge \Psi)$. □

The other lemma concerns when we can collapse multiple logical possibility operators into a single operator.

Lemma 13.6.2 (Diamond Collapsing)**.** *If* $\mathcal{L}' \supseteq \mathcal{L}$ *then*

$$\Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}\Phi \leftrightarrow \Diamond_{\mathcal{L}}\Phi$$

Proof. To prove the left to right direction, suppose that $\Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}\Phi$. Enter the $\Diamond_{\mathcal{L}}$ context. In this context we have $\Diamond_{\mathcal{L}'}\Phi$. Since $\mathcal{L}' \supset \mathcal{L}$, by Reducing we can infer $\Diamond_{\mathcal{L}}\Phi$. Exiting the $\Diamond_{\mathcal{L}}$ context, we have $\Diamond_{\mathcal{L}}(\Diamond_{\mathcal{L}}\Phi)$ in our original contest. So we can apply $\Diamond$ Elimination to infer $\Diamond_{\mathcal{L}}\Phi$.

1	$\Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}\Phi$	[1]
2	$\Diamond$ | $\Diamond_{\mathcal{L}'}\Phi$	1 In $\Diamond$ I [2*]
3	| $\Diamond_{\mathcal{L}}\Phi$	2 Reducing [2*]
4	$\Diamond_{\mathcal{L}}(\Diamond_{\mathcal{L}}\Phi)$	1,2-3 In $\Diamond$ E [1]
5	$\Diamond_{\mathcal{L}}\Phi$	4 $\Diamond$ Elimination [1]
6	$\Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}\Phi \rightarrow \Diamond_{\mathcal{L}}\Phi$	5 $\rightarrow$ I

To prove the other direction, suppose that $\Diamond_{\mathcal{L}}\Phi$. Entering this diamond context, we have Φ and can infer that $\Diamond_{\mathcal{L}'}\Phi$ by $\Diamond$ Introduction. So completing our inner diamond argument gives us $\Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}\Phi$.

1	$\Diamond_{\mathcal{L}}\Phi$	[1]
2	$\Diamond$ | Φ	[$\mathcal{L}$] 1 In $\Diamond$ I [2*]
3	| $\Diamond_{\mathcal{L}'}\Phi$	3, $\Diamond$ I [2*]
4	$\Diamond_{\mathcal{L}}(\Diamond_{\mathcal{L}'}\Phi)$	1, 2-3 In $\Diamond$ E [1]

□

We also observe that there is a corresponding $\Box$ version of the above lemma.

Lemma 13.6.3 (Box Collapsing). *If* $\mathcal{L}' \supseteq \mathcal{L}$ *then*

$$\Box_{\mathcal{L}}\Phi \leftrightarrow \Box_{\mathcal{L}}\Box_{\mathcal{L}'}\Phi$$

Proof. Note that this is equivalent to proving

$$\neg\Box_{\mathcal{L}}\Phi \leftrightarrow \neg\Box_{\mathcal{L}}\Box_{\mathcal{L}'}\Phi$$

which is just

$$\Diamond_{\mathcal{L}}\neg\Phi \leftrightarrow \Diamond_{\mathcal{L}}\Diamond_{\mathcal{L}'}\Phi$$

This is true by Diamond Collapsing.

$\Box$

Chapter 14

The Formal System II: Other Inference Rules

14.1 Cutback

Axiom 14.1.1 (Cutback)**.** If For any list of relations $\mathcal{L} = R_1, \ldots, R_m$,

$$[(\exists x)P(x) \wedge (\forall x \mid \text{Ext}(\mathcal{L})(x))\, P(x)] \rightarrow \Diamond_{\mathcal{L},P}(\forall x)P(x)$$

This axiom schema expresses the idea that if a predicate P applies to all the objects which relations in $\mathcal{L}$ apply to (and P applies to at least one thing), then it is logically possible to have a cut back universe which preserves how P and relations in $\mathcal{L}$ apply, and contains no objects outside the extension of P.

14.2 Modal Comprehension

Our next axiom schema, Modal Comprehension, expresses a somewhat similar idea to the Simple Comprehension axiom schema above. Modal Comprehension expands on the idea behind Simple Comprehension by ensuring the logical possibility of a relation which applies to exactly those objects picked out by some *modal* sentence. Informally, the idea is that Modal Comprehension lets us make inferences like the following

> SIBLINGS: Holding fixed the facts about the relations Married(x, y) and Sibling(x, y) it is logically possible to have a relation $R(x)$ that applies to exactly those married individuals x with more siblings than their spouse.

Note that having more siblings than one's spouse has to be cashed out in terms of the logical possibility of a surjective but not injective map from their siblings to those of their spouse. On first glance, it would appear this would require passing x (the individual for whom we wish to compare their siblings to those of their spouse) into the logical possibility operator evaluating the possibility of such a pairing. However, our language of logical possibility does not allow this kind of quantifying in.

Instead, we do this by using a special, otherwise-unused, n-place relation Q to label and preserve a choice for an n-tuple of objects in Ext$(\mathscr{L})$. We say that it is possible (fixing the $\mathscr{L}$ facts) for R to apply in such a way that, necessarily (fixing the $\mathscr{L}, R$ facts), R only relates objects in Ext$(\mathscr{L})$ and however Q chooses a unique n-tuple of objects in Ext$(\mathscr{L})$ for consideration,

R applies to this n-tuple iff a certain modal claim ϕ describing the behavior of $\mathscr{L}$ and Q is true. In this case, the relevant $\mathscr{L}$ is $\{\text{Married}, \text{Sibling}\}$, and the modal sentence ϕ is $\Diamond_{\text{Married},\text{Sibling},Q}$ $((\exists x)Q(x) \wedge (\exists y)\text{Married}(x, y)$ and $Z(\cdot, \cdot)$ is a surjective but not injective map from the siblings of x to those of y.

We can thus express the informal claims like siblings with a sentence of the following form:

$$\Diamond_{\text{Married},\text{Sibling},Q}\Box_{\text{Married},\text{Sibling},\mathscr{Q},R}(\exists! x \mid Q(x)) \rightarrow (\exists x)(Q(x) \wedge [R(x) \leftrightarrow x \in \text{Ext}(\mathscr{L}) \wedge \phi)] \tag{14.1}$$

Since it is possible for Q to apply to any single object x, the necessity operator above ensures that R applies to exactly those x which have more siblings than their spouse. Or, to put the point differently, if there was some x in the extension of one of $\text{Married}, \text{Sibling}, Q$ where $R(x) \leftrightarrow \phi$ then, intuitively, it would be possible for Q to apply to such an x contradicting the assumption. With this motivation in place, I can now state the Modal Comprehension Schema as follows.

Axiom 14.2.1 (Modal Comprehension)**.** If

- R does not occur in $\mathcal{L}, \Psi$ or ϕ
- Q does not occur in $\mathcal{L}$ or Ψ
- ϕ is content restricted to $\mathcal{L}, Q$

then

$$\Psi \to \Diamond_{\mathcal{L}}\left(\Psi \wedge \Box_{\mathcal{L},R}\left[(\exists! \vec{x} \mid Q(\vec{x})) \to (\exists \vec{x} \mid Q(\vec{x}))\left[R(\vec{x}) \leftrightarrow \text{Ext}(\mathscr{L})(\vec{x}) \wedge \phi(\vec{x}\right]\right]\right)$$

where $(\exists! \vec{x} \mid Q(\vec{x}))$ means there is a unique tuple $\vec{x}$ satisfying Q. Note that we take $(\exists! \vec{x} \mid Q(\vec{x}))[\varphi(\vec{x})]$ to mean that Q applies to a unique n-tuple of objects $\vec{x}$ and those objects satisfy (not necessarily uniquely) $\varphi(\vec{x})$.

14.3 Possible Infinity

We now adopt an axiom asserting the logical possibility of infinitely many objects.

Axiom 14.3.1 (Infinity)**.** $\Diamond\Psi$ where Ψ is the conjunction of the following claims:

1. The successor of an object is unique $(\forall x)(\forall y)(\forall y')\,[S(x,y) \wedge S(x,y') \to y = y']$
2. successor is one-to-one $(\forall x)(\forall y)(\forall x')(S(x,y) \wedge S(x',y) \to x = x')$
3. there is a unique object that has a successor and isn't the successor of anything $(\exists! x : (\exists y)S(x,y) \wedge (\forall y)\neg S(y,x))$
4. everything that is a successor has a successor $(\forall x)[(\exists y)S(y,x) \to (\exists z)S(x,z)]$
5. S is anti-reflexive: $(\forall x)(\forall y)\,[S(x,y) \to \neg S(y,x)]$

Note that by 1 the relation S is a function.

14.4 Possible Powerset

Axiom 14.4.1 (Possible Powerset)**.** If F, C are distinct predicates, $\in_C$ a two-place relation, then $\Diamond_F \mathscr{C}(C, \in_C, F)$.

Here $\mathscr{C}(C, \in_C, F)$ is the conjunction of the following claims:

- $(\forall x)\neg\,(C(x) \wedge F(x))$, i.e., the objects satisfying F and C are disjoint.

- $(\forall x)(\forall y)\,(x \in_C y \to F(x) \wedge C(y))$.

- $\Box_{C,\in_C,F}(\exists x)\,[C(x) \wedge (\forall y)\,((F(y) \wedge K(y)) \leftrightarrow y \in_C x)]$, i.e., it's necessary that however some predicate K applies to some objects satisfying F, there exists a corresponding class C whose elements are exactly the objects which F applies to.

- $(\forall y)(\forall y')(C(y) \wedge C(y') \wedge \neg y = y' \to (\exists x)\neg(x \in_C y \leftrightarrow x \in_C y')$, i.e., classes are extensional (no two members of C contain, in the sense of $\in_C$, the same elements).

Intuitively, this axiom schema says that it is always possible to add a layer of classes to the objects satisfying some predicate F. Note that $\mathscr{C}(C, \in_C, F)$ is content restricted to $C, \in_C, F$.

We now prove a simple lemma that we will often use in applying this axiom. Since this strictly strengthens the above axiom, we will take invocations of possible powerset to refer to applications of the following lemma.

Lemma 14.4.2 (Possible Powerset)**.** *Suppose that $\mathcal{L}' \supset \mathcal{L}$ and neither C*

nor $\in_C$ *are in* $\mathcal{L}'$ *then*

$$\Diamond_{\mathcal{L}'}\mathscr{C}(C, \in_C, \mathrm{Ext}(\mathcal{L}))$$

Proof. We note that it is enough to prove

$$\Diamond_{\mathcal{L}'}(\forall z)(F(x) \leftrightarrow \mathrm{Ext}(\mathcal{L})(x)) \wedge \mathscr{C}(C, \in_C, F)$$

for some new predicate F not in $\mathcal{L}'$ since the formula inside the $\Diamond_{\mathcal{L}'}$ is first order equivalent to the desired claim (hence the claimed result follows by Logical Closure).

By Simple Comprehension we have

$$\Diamond_{\mathcal{L}'}(\forall x)\,[F(x) \leftrightarrow \mathrm{Ext}(\mathscr{L})(x)]$$

Enter this $\Diamond_{\mathcal{L}'}$ context, i.e., use Inner Diamond to reason about what else must be true in this context, and invoke Possible Powerset to derive $\Diamond_F(\mathscr{C}(C, \in_C, F))$. Via $\Diamond$ Ignoring we can infer $\Diamond_{\mathcal{L}',F}(\mathscr{C}(C, \in_C, F))$. By Importing and Logical Closure we can infer

$$\Diamond_{\mathcal{L}',F}(\forall z)(F(x) \leftrightarrow \mathrm{Ext}(\mathcal{L})(x)) \wedge \mathscr{C}(C, \in_C, F)$$

The desired conclusion follows by leaving the $\Diamond_{\mathcal{L}'}$ context and applying Diamond Collapsing.

□

14.5 Choice

Axiom 14.5.1 (Choice)**.** For all $n \geq 0, m > 0$ if I is an n-ary relation (where a 0-ary relation is assumed to be $\neg\bot$) and $R, \widehat{R}$ are $n+m$-ary relations with $\widehat{R}$ not appearing in Φ or $\mathcal{L}$ (nor equal to I, R) then

$$\Phi \wedge (\forall \vec{x})\Big[I(\vec{x}) \rightarrow (\exists \vec{y})R(\vec{x},\vec{y})\Big] \rightarrow$$
$$\Diamond_{\mathcal{L},I,R}\Phi \wedge \left[(\forall \vec{x})(\forall \vec{y})\Big(\widehat{R}(\vec{x},\vec{y}) \rightarrow R(\vec{x},\vec{y})\Big) \wedge (\forall \vec{x})\Big[I(\vec{x}) \rightarrow (\exists! \vec{y})\widehat{R}(\vec{x},\vec{y})\Big]\right]$$

This axiom schema captures the same intuition as the axiom of choice in set theory. It says that if every x satisfying I is related to some y by R, then (fixing I, R) another relation $\widehat{R}$ can behave like a choice function selecting a unique such y for each x.

Note that in the case where $n = 0$ (when I becomes just a necessary truth) the axiom asserts the possibility of an $\widehat{R}$ which applies to a unique m-tuple in the extension of R. The utility of this special case is demonstrated in the following lemma[1].

Proposition 14.5.2 (Simplified Choice)**.** *Suppose $\widehat{R} \notin \mathcal{L}, R \in \mathcal{L}$ (and $\widehat{R}$ is not the same relation as R).*

$$(\exists \vec{x})\,(R(\vec{x})) \wedge \Phi \rightarrow \Diamond_{\mathcal{L}}\left(\Phi \wedge (\exists \vec{x})\left[\widehat{R}(\vec{x}) \wedge \widehat{R}(\vec{x}) \wedge (\forall \vec{y})\,(P(\vec{y}) \rightarrow \vec{x} = \vec{y})\right]\right)$$

Proof. This follows directly from Choice by letting $n = 0$ as we regard the

[1] Note that in actualist set theory this fact is guaranteed by applying comprehension with parameters but our version of comprehension doesn't allow this.

0-ary relation I appearing in Choice as $\neg\bot$. □

14.6 Possible Amalgamation

My last, and most complex, axiom schema articulates a modal version of the set theorist's Axiom of Replacement. This axiom captures the intuition that if the most general laws of logic permit a certain scenario $\phi(x)$ for each x in some logically possible I then they don't forbid the 'disjoint union' of these scenarios.

Crudely speaking, this axiom takes us from the logical possibility (given some starting structure $\mathcal{L}$), of satisfying a certain formula $\phi(x)$ *for any single* x in a base collection of objects (those satisfying some I in $\mathcal{L}$), to the logical possibility of an expanded universe where *for every* object x satisfying I, there is a corresponding structure (indexed to this object x) within which a version of $\phi(x)$ is true.

For example, if we take I to be the predicate person$(\cdot)$ and the $\mathcal{L}$ to be the list person$(\cdot)$, childOf(x, y) Possible Amalgamation licenses claims like:

> If, for any choice of a person, there could be (holding fixed the facts about people and parentage) as many ghosts as that person has children, then (holding fixed the facts about people and parentage) it could be that, for every person x, there are as many ghosts-haunting-x (disjoint from everyone else's ghosts) as x has children.

While this principle in some sense serves the same purpose as the set theorist's axiom of replacement it differs in a critical way from actualist Axiom of Replacement. Actualist replacement acts as a closure condition on a single structure (the hierarchy of sets) and it's this aspect which makes its consistency non-obvious. However, the Amalgamation axiom merely asserts the *possibility* of a scenario in which this disjoint union is realized. This possibility is obviously logically possible in a way that assuming the existance of a single structure closed under replacement is not[2].

As before, articulating this principle can seem to require quantifying in to the $\Diamond$ of logical possibility. However, we can use the same trick (involving an otherwise unused predicate Q) to get around it, as we did when formulating modal comprehension above.

Axiom 14.6.1 (Amalgamation)**.** If

- $\mathcal{L}$ is a list of relations which contains the predicate I but not Q or $R_1 ... R_n$
- Φ is content-restricted $\mathcal{L}, Q, R_1 \ldots R_n$. (where P, $R_1 \ldots R_n$ and $\mathcal{L}$ share no relations)
- $\widehat{R_1} \ldots \widehat{R_n}$ are otherwise unused relations such that if R_i is an n-place relation $\widehat{R}_i$ is an $n+1$ place relation.

Let $\Psi(x)$ be the formula

[2]While earlier potentialist approaches, e.g., Hellman [44], did try to justify their uses of replacement I believe this axiom is more clearly guaranteed not to contravene the most general laws of logic.

$$\bigwedge_{1\leq i\leq n} (\forall \vec{v})(R_i(\vec{v}) \leftrightarrow \widehat{R_i}(\vec{v}, x))$$

asserting that $\widehat{R}_i$ with x inserted into the last place behaves exactly the same as R_i. Let $\pi(x, y)$ be the formula

$$\bigvee_{\substack{1\leq i\leq n\\ 1\leq j\leq l_i}} (\exists z_1)\ldots,(\exists z_{j-1}),(\exists z_{j+1}),\ldots,(\exists z_{l_i})\widehat{R}_i(z_1,\ldots,z_{j-1},x,z_{j+1},\ldots,z_{l_i},y)$$

which asserts that x appears in some tuple ending with y satisfying some $\widehat{R}_i$

then

$$\begin{aligned}
&\Box_{\mathcal{L}}\left[(\exists! x \mid Q(x))(I(x)) \rightarrow \Diamond_{\mathcal{L},Q}\Phi\right] \rightarrow \\
&\Diamond_{\mathcal{L}}\Big[(\forall x)(\forall y)(\forall y')\Big[(\neg y = y' \wedge \pi(x,y) \wedge \pi(x,y') \rightarrow x \in \mathrm{Ext}(\mathcal{L})\Big] \wedge \\
&\qquad \Box_{\mathcal{L},\widehat{R_1}\ldots\widehat{R_n}}\left[(\exists! x)(Q(x))\left(Q(x) \wedge I(x) \wedge \Psi(x)\right) \rightarrow \Phi\right]\Big]
\end{aligned}$$

Remember that $(\exists! x \mid Q(x))(I(x))$ indicates that there is a unique x such that Q and that this x also satisfies I.

To see how this captures the intuition from the start of the section, note that (informally) the antecedent merely asserts that for each x satisfying I (with the value of x conveyed by Q) it is logically possible that $\Phi(x)$. The consequent, in turn, asserts that it's logically possible to have a single logically possible structure which can be broken up into disjoint 'domains' M_x for each

x satisfying I such that $M_x \models \Phi$. Note that M_x here is just a way of talking about the objects satisfying $\widehat{R}_i(\cdot, \ldots, \cdot, x)$ and we capture talk of modelling by considering the logically possible scenario in which $R_i(y_1, \ldots, y_n)$ holds iff $\widehat{R}_i(y_1, \ldots, y_n, x)$. In other words this complex sentences merely expresses the relatively straightforward intuition that if we can index logically possible scenarios by I then it's logically possible to have a disjoint union of scenarios witnessing all the logical possibility facts indexed by I.