Melvin Fitting

Intensional Logic — Beyond First Order

Abstract. Classical first-order logic can be extended in two different ways to serve as a foundation for mathematics: introduce higher orders, type theory, or introduce sets. As it happens, both approaches have natural analogs for quantified modal logics, both approaches date from the 1960's, one is not very well-known, and the other is well-known as something else. I will present the basic semantic ideas of both higher order intensional logic, and intensional set theory. Before doing so, I'll quickly sketch some necessary background material from quantified modal logic. Except for standard material concerning propositional modal logics, the paper is essentially self-contained.

Keywords: intensional logic, modal logic, set theory, type theory, higher order logic, forcing

1. Introduction

A little way into the twentieth century, both Bertrand Russell and Ernst Zermelo discovered the now well-known paradox about the set of sets not having themselves as a member (or of the predicate that can be predicated of those predicates not predicatable of themselves). Each came up with a solution, and while there are similarities, the differences are striking. Russell, [Rus08], created type theory, which separates predicates into distinct categories—types—each predicate falling into exactly one category. The types themselves have a rather labyrinthine structure, but at least that structure is a well-founded one. That is, if one starts with a predicate of arbitrary type, moves to a predicate of the type of one of its arguments, then to a predicate of the type of an argument of that, and so on, one must eventually 'bottom out.' Zermelo, taking a different route, introduced the cumulative hierarchy of sets, with the universe of sets divided into levels indexed by ordinals, [Zer08, Zer35]. It is more complicated than type theory, in a way, because the levels extend through the entire class of ordinals. On the other hand, it is simpler in that it is cumulative: a set present at any level is also present at all higher ones. This means that, unlike in Russell's theory, we do not find a distinct copy of the natural numbers at each type level—one copy serves for all. Still, as in type theory, a set in Zermelo's hier-

Presented by **Name of Editor**; *Received* December 1, 2002

Studia Logica **68**: 1–**??**, 2001.

archy can only have as members things of lower levels, and this (apparently) serves to avoid paradox.

Both type theory and the cumulative hierarchy of sets were introduced to provide a safe foundation for mathematics, and both apparently do that. But since their introduction, the two have gone their separate ways. The cumulative hierarchy of sets allows mathematics to be developed in a simpler and, most feel, more natural way, so type theory has largely been ignored by mathematicians. On the other hand, the constructs of language often seem to have a type structure to them. We say of a person that they are good, or tall (or not). We say of goodness that it is a virtue while tallness is not. We do not say a person is a virtue or that virtue is tall. It is not simply that saying these would be to utter falsehoods—we do not say them because the subject/predicate combinations are inappropriate—saying them results in an ill-typed sentence. Since type theory seems to have reasonable relationships to natural language, it should not come as a surprise to find that it has found important applications in computer science, in the design of programming languages. Imposing a type structure has been found to lead to languages in which certain kinds of programming errors are much harder to make—they become grammatical errors which are easily found by a compiler. Even a language like Lisp, which is untyped, is often used with an implicit typing in mind by the programmer.

The history of modal logic has often paralleled that of classical logic, but with a time delay of a generation or two. Propositional modal logics are now standard tools. First-order modal logics have a remarkably rich structure that is currently being explored by researchers. But what about the analogs of the work of Russell and Zermelo? As it happens, both have modal—intensional—versions, and both are in the literature, though their appearance followed the classical versions by perhaps half a century. Even so, they are not well-known. Indeed, the intensional version of Zermelo's cumulative set hierarchy is in the peculiar position of being well-known, but as something else. It is my aim, in this paper, to sketch the basic ideas. What we will be looking at should be standard tools, and I hope to move them a little in that direction.

Since both intensional type theory and intensional set theory are complex organisms, I think it best to lead up to them gradually, so that not all the essential ideas turn up in a bunch. Consequently I begin with brief sketches of simpler structures involving quantification. These are certainly more familiar to readers. If they are too familiar, I apologize in advance.

2. First-Order Background

I assume everyone reading this paper is thoroughly familiar with propositional modal logics, but there are some points concerning first-order versions that may be less well-known, so a little background discussion of these issues is appropriate. I am only interested in semantics here—axiom systems or other proof procedures are ignored in this discussion. Also, there are two broadly differing approaches to quantified modal semantics, one using counterpart relations, the other not. I follow the tradition, dating from Kripke, that does not use counterpart relations—when one talks of Niagara Falls under other possible circumstances, one does not mean a counterpart of Niagara Falls, one means Niagara Falls. However, for those wishing to read further on the counterpart version, I recommend [Lew68, Lew71] for the early work, and [Kra93, KK01, KK03, Kut02], and the references given in these papers, for recent work. Also, some discussion of relationships between the two approaches can be found in [Fit01].

To keep things relatively simple, throughout this paper I'll assume languages have no constant or function symbols. Formulas are built up from atomic formulas involving relation symbols and variables, using propositional connectives, quantifiers, and modal operators. I'll take some of these as primitive and the others as defined, as convenient.

2.1. Rigid Objects

Perhaps the simplest first-order modal assumption in the literature is that quantifiers range over objects that are *rigid*, the same from world to world, and that such objects exist at all worlds. This is the so-called *constant domain* semantics. Let's begin here.

Definition 2.1. The structure $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{D}, \mathcal{I} \rangle$ is a *constant domain* first-order model if $\mathcal{G}$ is a non-empty set of possible worlds, $\mathcal{R}$ is an accessibility relation on $\mathcal{G}$, $\mathcal{D}$ is a non-empty domain, and $\mathcal{I}$ is an interpretation, assigning to each n-place relation symbol a mapping from possible worlds to n-place relations on $\mathcal{D}$. A *valuation* in $\mathcal{M}$ is a mapping from variables to $\mathcal{D}$. Valuations are not world-dependent, unlike interpretations.

$\mathcal{M}, \Gamma \Vdash_v X$ is intended to symbolize that formula X is true at world Γ of model $\mathcal{M}$ with respect to valuation v. It is characterized as follows.

1. For atomic formulas, $\mathcal{M}, \Gamma \Vdash_v P(x_1, \ldots, x_n)$ if $\langle v(x_1), \ldots, v(x_n) \rangle \in \mathcal{I}(P)(\Gamma)$
2. $\mathcal{M}, \Gamma \Vdash_v X \wedge Y$ iff $\mathcal{M}, \Gamma \Vdash_v X$ and $\mathcal{M}, \Gamma \Vdash_v Y$

3. $\mathcal{M}, \Gamma \Vdash_v \neg X$ iff not-$\mathcal{M}, \Gamma \Vdash_v X$

4. $\mathcal{M}, \Gamma \Vdash_v (\forall x)\Phi$ iff $\mathcal{M}, \Gamma \Vdash_w \Phi$ for every valuation w that differs from v on at most x

5. $\mathcal{M}, \Gamma \Vdash_v \Box X$ iff $\mathcal{M}, \Delta \Vdash_v X$ for every $\Delta \in \mathcal{G}$ such that $\Gamma \mathcal{R} \Delta$

This is a well-known modal semantics—perhaps the best known of those for quantified modal logic—so I need say little more about it here. The next most complicated after the constant domain semantics is the varying domain version, which still assumes rigidity of objects. Since the semantics is similar to that of Definition 2.1, I will only say wherein the differences lie.

Definition 2.2. $\langle \mathcal{G}, \mathcal{R}, \mathcal{D}, \mathcal{I} \rangle$ is a *varying domain* first order model if it meets the conditions of Definition 2.1, except that $\mathcal{D}$ is a *mapping* that assigns to each member of $\mathcal{G}$ some non-empty set (the domain at that possible world), and $\mathcal{I}$ assigns to each n-place relation symbol a mapping from possible worlds to n-place relations on $\cup_{\Gamma \in \mathcal{G}} \mathcal{D}(\Gamma)$.

The definition of truth at a world is also as in Definition 2.1, except for the following.

4. $\mathcal{M}, \Gamma \Vdash_v (\forall x)\Phi$ iff $\mathcal{M}, \Gamma \Vdash_w \Phi$ for every valuation w that differs from v on at most x, *and is such that* $w(x) \in \mathcal{D}(\Gamma)$

With varying domain models, quantification at a world is over the domain associated with that world, and not over a common world-independent domain. One can think of the domain associated with a world as the set of things *actually* existing at that world, and so the semantics is sometimes called *actualist*. In a similar way, constant domain semantics is sometimes called *possibilist*—one thinks of the common domain as what does, or what might exist, where a possible existent is an actual existent in some possible world. Understanding the differences between constant and varying domain models, and their connections with the Barcan formula and its converse, was an early achievement of first-order modal logic.

As it happens, the possibilist semantics, which is the simpler one technically, can serve very well for both possibilist and actualist purposes, a point discussed in both [HC96] and in [FM98]. Suppose we have a varying domain model, $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{D}, \mathcal{I} \rangle$. Using it, define a new constant domain $\mathcal{D}' = \cup_{\Gamma \in \mathcal{G}} \mathcal{D}(\Gamma)$. Also, introduce a new one-place relation symbol E to play the role of an *existence* predicate, and let $\mathcal{I}'$ be like $\mathcal{I}$ except that it is extended to E by setting $\mathcal{I}'(\mathsf{E})$ to be the mapping sending each possible world Γ to $\mathcal{D}(\Gamma)$. Now we have a constant domain model, $\mathcal{M}' = \langle \mathcal{G}, \mathcal{R}, \mathcal{D}', \mathcal{I}' \rangle$.

Suppose we translate each formula X not containing E to a formula X^{E} by relativizing each quantifier: replace $(\forall x)\Phi$ by $(\forall x)[\mathsf{E}(x) \supset \Phi]$ and $(\exists x)\Phi$ by $(\exists x)[\mathsf{E}(x) \wedge \Phi]$. It is not hard to show that $\mathcal{M}, \Gamma \Vdash_v X$ iff $\mathcal{M}', \Gamma \Vdash_v X^{\mathsf{E}}$. In other words, varying domain semantics can be embedded into constant domain semantics provided we give up the notion that existence must be captured by the existential quantifier.

From now on in this paper, I will only consider constant domain models. As just noted, this is no real restriction, though it certainly simplifies things. But this still leaves the question of whether the assumption of rigidity is an appropriate one, and this point requires a section to itself.

2.2. Non-Rigid Objects

The objects that quantifiers range over in the previous section are *rigid*—they are like the Eiffel tower, or George Washington, the same thing from possible world to possible world. But *non-rigid* objects are also quite reasonable to work with—the gross domestic product of Denmark, or the Secretary-General of the United Nations, say—varying in designation from circumstance to circumstance. It is clear how to model such things semantically: use functions from possible worlds to rigid objects. But complexities force themselves on us. For instance, if we say the gross domestic product of Denmark currently is $\$171.6 \times 10^9$ (true at the time of writing), we are relating something that is non-rigid to something that is rigid. The need to talk about rigid and non-rigid items together means we need a more complex notion of atomic formula.

For the rest of this section I'll assume that each relation symbol has a *type* associated with it, where a type is an n-tuple whose entries are either I or O. Informally, O is for *object*, the kind of rigid thing we considered in the previous section, while I stands for *intension*, which we intend to be something non-rigid. For example, a type of $\langle I, O\rangle$ would be appropriate for saying the gross domestic product of Denmark is $\$171.6 \times 10^9$ We also need two kinds of variables, call them intensional, or type I, and extensional, or type O. If R is a relation symbol of type $\langle T_1, \ldots, T_n\rangle$, where each T_i is either I or O, then for $R(x_1, \ldots, x_n)$ to be an atomic formula, x_i must be an intensional variable if T_i is I, and extensional if T_i is O. All this is rather straightforward.

If f is an intension, sometimes we want to talk about f itself, and sometimes about what f designates in a particular context. For instance, if we say "the gross domestic product of Denmark is a useful economic measure," we refer to the intension, gross-domestic-product-of-Denmark.

But if we say "the gross domestic product of Denmark is currently greater than the gross domestic product of Finland," we refer to the amount of value gross-domestic-product-of-Denmark designates at the present moment, $\$171.6 \times 10^9$. It is useful to have notation to distinguish these sorts of things, and I introduce an *extension of* operator, $\downarrow$, for this purpose. If x is an intensional variable, $\downarrow x$ is extensional, while $\downarrow$ is not applicable to extensional variables. The idea is, the value associated with $\downarrow x$ in a context is to be whatever extensional object is designated by the intension associated with x in that context. It is allowed that $\downarrow x$ can fill an O position in an atomic formula, if x can fill an I one. I'll call $\downarrow x$ *relativized*, since its semantical behavior will depend on context. This semantic behavior will be specified below, and it will be clear that it corresponds to the informal ideas just discussed.

As it happens, on combining the machinery discussed so far, we find ambiguities. If we say the Secretary-General of the United Nations cannot be younger than 65, in a sense this is true, because at the time of writing this paper the Secretary-General is Kofi Annan, and he is already 65. But we can also say the Secretary-General of the United Nations might be younger than 65, and this too is plausible because the Secretary-Generalship probably will be held by a younger person someday. Suppose we informally think of P as a relation symbol of type $\langle O\rangle$ intended to say, of its argument, that its age is less than 65. Now, our problem is that if x is an intensional variable whose intended value is the intension Secretary-General-of-the-United-Nations, our first informal assertion above would be formalized as $\neg\Diamond P(\downarrow x)$ while the second becomes $\Diamond P(\downarrow x)$. Since both are plausible, we clearly have a problem. Of course, the problem arises because the action of letting x designate, that is, of evaluating $\downarrow x$, and the action of passing to an alternative possible world, that is, of interpreting $\Diamond$, are not actions that commute.

To disambiguate this, one more piece of machinery is needed. It will play a much bigger role later on when the full type hierarchy is introduced, but as we have just seen, it is needed now too.

DEFINITION 2.3. If $x_1, \ldots, x_n$ is a sequence of variables, of respective types $T_1, \ldots, T_n$, and Φ is a formula, then $\langle\lambda x_1, \ldots, x_n.\Phi\rangle$ is a *predicate abstract*, of type $\langle T_1, \ldots, T_n\rangle$. It can be used like a relation symbol in forming atomic formulas, that is, if $t_1, \ldots, t_n$ are of types $T_1, \ldots, T_n$ respectively, then $\langle\lambda x_1, \ldots, x_n.\Phi\rangle(t_1, \ldots, t_n)$ is an atomic formula.

With this extension of syntax, the problematic formula $\Diamond P(\downarrow x)$ splits into $\langle\lambda y.\Diamond P(y)\rangle(\downarrow x)$ and $\Diamond\langle\lambda y.P(y)\rangle(\downarrow x)$, where x is an intensional variable whose value is intended to be the Secretary-General-of-the-United-Nations intension. Of course we need a semantics that gives these formulas different behaviors. The following is taken from [Fit01], where all this was introduced under the name **FOIL**, for *first-order intensional logic*. A very full discussion of the essential ideas involved can be found in [FM98]. Note that, following up on earlier discussion, domains of quantification are constant, not varying.

Definition 2.4. A **FOIL** model is a structure $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{D}_O, \mathcal{D}_I, \mathcal{I} \rangle$ where: $\mathcal{G}$ and $\mathcal{R}$ are as usual, $\mathcal{D}_O$ is a non-empty set, the *object domain*, $\mathcal{D}_I$ is a non-empty set of functions from $\mathcal{G}$ to $\mathcal{D}_O$, the *intension domain*, and $\mathcal{I}$ is an *interpretation* such that, if P is a relation symbol of type $\langle T_1, \ldots, T_n \rangle$ then $\mathcal{I}(P)$ is a mapping from $\mathcal{G}$ to subsets of $\mathcal{D}_{T_1} \times \cdots \times \mathcal{D}_{T_n}$.

A *valuation* in $\mathcal{M}$ is a map that assigns to each type O variable a member of $\mathcal{D}_O$ and to each type I variable a member of $\mathcal{D}_I$. If v is a valuation, let $\hat{v}$ be a mapping on variables *and relativized variables* to functions on worlds, given by: if x is unrelativized, $\hat{v}(x)$ is the constant map assigning to each world Γ the entity $v(x)$, and if $\downarrow x$ is relativized, $\hat{v}(\downarrow x)$ is the map assigning to world Γ the object $v(x)(\Gamma)$.

Truth at a world is defined more or less as expected. There are two kinds of quantifiers, since there are two kinds of variables—if an object variable is involved, quantification is over $\mathcal{D}_O$, and if an intension variable is involved, quantification is over $\mathcal{D}_I$. The main point is that condition 1 of Definition 2.1 is replaced with the following.

1. For atomic formulas,

   (a) if $P(x_1, \ldots, x_n)$ is an atomic formula (so $x_1, \ldots, x_n$ are extensional variables, intensional variables, and relativized intensional variables), $\mathcal{M}, \Gamma \Vdash_v P(x_1, \ldots, x_n)$ iff $\langle \hat{v}(x_1)(\Gamma), \ldots, \hat{v}(x_n)(\Gamma) \rangle \in \mathcal{I}(P)(\Gamma)$

   (b) $\mathcal{M}, \Gamma \Vdash_v \langle \lambda x_1, \ldots, x_n.\Phi \rangle(t_1, \ldots, t_n)$ iff $\mathcal{M}, \Gamma \Vdash_w \Phi$, where w is like v except that $w(x_i) = \hat{v}(t_i)(\Gamma)$, for $i = 1, \ldots, n$.

In **FOIL**, the use of predicate abstraction is essentially as a scope indicator. In $\langle \lambda y.\Diamond P(y) \rangle(\downarrow x)$ the x is being used with broad scope while in $\Diamond \langle \lambda y.P(y) \rangle(\downarrow x)$ it is used with narrow scope. These notions essentially trace back to Russell's work on definite descriptions, [Rus05], but are also significant for intensional issues. The logic **FOIL** is a flexible and natural logic, by design with both extensional and intensional features, but here it is simply presented as a lead-in to what follows, essentially to introduce some notation and basic ideas.

3. Classical type theory

Russell intended to base the foundations of mathematics purely on classical logic. The number 3, for instance, was to be the property that applied to just those properties that applied to three items. But this required that properties be applicable to other properties, and this circumstance eventually lead to paradox. After a variety of other attempts, Russell's eventual way of dealing with paradox was to introduce type theory. The first version, [Rus08], and in *Principia Mathematica*, [WR27], involved *ramified* types, but eventually this evolved into what has become known as the *simple theory of types*. This has been given elegant formulations in more than one way—here I'll quickly sketch one version, as further lead-in to the modal system presented in the next section. The basic idea is to divide the universe of discourse into separate categories, types, with a predicate of one type applicable only to items of certain other designated types. But unlike the baby notion of types from the previous section, now we need predicates, predicates of predicates, predicates of these, and so on. The type structure is intended to avoid any hazardous circularity of reference, thus avoiding paradoxes, while still providing a rich enough structure within which to carry out the development of mathematics.

Definition 3.1. 0 is a type. If $T_1, \ldots, T_n$ are types, $\langle T_1, \ldots, T_n \rangle$ is a type.

The idea is, 0 is intended to be the type of ground-level objects, and $\langle T_1, \ldots, T_n \rangle$ is intended to be the type of a predicate that takes n arguments, of types $T_1, \ldots, T_n$ respectively.

For each type T I'll assume there are infinitely many variables of that type. The type of a variable will be clear from context, or else I'll simply say in words what the type is. Formulas come next, but their definition is more complex then in a first-order setting.

Definition 3.2. Formulas, terms, and predicate abstracts are characterized (simultaneously) as follows.

1. Let Φ be a formula and $x_1, \ldots, x_n$ be a sequence of distinct variables of types $T_1, \ldots, T_n$ respectively. $\langle \lambda x_1, \ldots, x_n.\Phi \rangle$ is a *predicate abstract*. Its type is $\langle T_1, \ldots, T_n \rangle$.

2. *Terms* are either predicate abstracts or variables. (Recall, I'm not considering constant or function symbols in this paper.)

3. If t is a term of type $\langle T_1, \ldots, T_n \rangle$, and $t_1, \ldots, t_n$ is a sequence of terms of types $T_1, \ldots, T_n$ respectively, then $t(t_1, \ldots, t_n)$ is an atomic formula, hence also a formula.

4. If Φ is a formula, so is $\neg\Phi$.

5. If Φ and Ψ are formulas so is $(\Phi \wedge \Psi)$.

6. If Φ is a formula and x is a variable then $(\forall x)\Phi$ is a formula.

Now that we have formulas, the next thing is to have a standard version of a model. In first-order logic there are relation symbols, and so an interpretation function is needed to give them semantic meanings. In the present setup there are no relation symbols as such—there are only variables, and predicate abstracts of non-ground types—so an interpretation function is not needed.

Definition 3.3. Let $\mathcal{D}$ be a non-empty set. For each type T a collection $[\![T, \mathcal{D}]\!]$ is defined as follows.

1. $[\![0, \mathcal{D}]\!] = \mathcal{D}$.

2. $[\![\langle T_1, \ldots, T_n \rangle, \mathcal{D}]\!]$ is the collection of all subsets of $[\![T_1, \mathcal{D}]\!] \times \cdots \times [\![T_n, \mathcal{D}]\!]$.

A member of $[\![T, \mathcal{D}]\!]$ is called an *object of type T over $\mathcal{D}$*. A *valuation* in $\mathcal{D}$ is a mapping v that assigns to each variable x of type T some object of type T over $\mathcal{D}$.

Now we can simultaneously define the notion of truth of a formula over a domain $\mathcal{D}$, and the denotation of a term, all with respect to a valuation. Just as in the previous section, I'll use $\hat{v}$ for the extension of v to more complex things than just variables—in this case, to predicate abstracts. Think of $\hat{v}$ as a mapping from terms to denotations.

Definition 3.4. Let $\mathcal{D}$ be a non-empty set, and let v be a valuation in it.

1. If x is a variable of type T, $\hat{v}(x) = v(x)$, and so $\hat{v}(x) \in [\![T, \mathcal{D}]\!]$.

2. If $\langle \lambda x_1, \ldots, x_n.\Phi \rangle$ is a predicate abstract of type T then $\hat{v}(\langle \lambda x_1, \ldots, x_n.\Phi \rangle)$ is the following member of $[\![T, \mathcal{D}]\!]$:

$$\{\langle w(x_1), \ldots, w(x_n) \rangle \mid \text{valuation } w \text{ agrees with } v \text{ on all variables except possibly } x_1, \ldots, x_n \text{ and } \Phi \text{ is true over } \mathcal{D} \text{ with respect to } w\}$$

3. An atomic formula $t(t_1, \ldots, t_n)$ is true over $\mathcal{D}$ with respect to v, if $\langle \hat{v}(t_1), \ldots, \hat{v}(t_n) \rangle \in \hat{v}(t)$.

4. $\neg\Phi$ is true over $\mathcal{D}$ with respect to v if Φ is not true over $\mathcal{D}$ with respect to v.

5. $\Phi \wedge \Psi$ is true over $\mathcal{D}$ with respect to v if both Φ and Ψ are true over $\mathcal{D}$ with respect to v.

6. $(\forall x)\Phi$ is true over $\mathcal{D}$ with respect to v if Φ is true over $\mathcal{D}$ with respect to every valuation w that agrees with v on all variables except possibly for x.

This is the natural and obvious semantics for the classical typed language that was introduced above. It is well known that validity with respect to this semantics does not correspond to any formal proof procedure. One needs a more general semantics—Henkin models—but these details are more than I want to get into here. Suffice it to say that the result is, in fact, suitable for Russell's purpose. Unfortunately, the result is not appealing to mathematicians. For instance, if T is a type, we can characterize the predicates of type $\langle T \rangle$ that apply to exactly three things of type T, say we call them holding-of-three-things predicates. Then we can introduce a predicate of type $\langle\langle T \rangle\rangle$ that applies to exactly the holding-of-three-things predicates of type $\langle T \rangle$, and we can identify this type $\langle\langle T \rangle\rangle$ predicate with the number 3, but note that it is the number 3 appropriate for type T objects. We actually get multiple copies of the number 3, one for each type T. This kind of thing is unfortunate, and doomed type theory as a mathematical tool, though as I noted earlier, it has found applications in linguistics and computer science.

This is the last of the sections intended to lead up to two specific intensional systems, about to be presented. I will say no more about the type system of this section, since it is subsumed in the system of the next section.

4. Higher-Order Modal Logic

While types present difficulties for mathematical applications, they are natural in other areas, in particular, in the formal treatment of natural language. Since modal constructs come up often in natural language expressions—temporal constructs especially—a combination of typed logic and Kripke structures is a natural thing to consider. In fact, precisely this combination has a venerable history—[Mon60, Mon68, Mon70, Gal75]. Here I present a variation on that work, amounting to a combination of type machinery from Section 3 with intensional machinery from Section 2.2. After the formal machinery has been introduced, I'll look at a few examples.

4.1. The Formal Machinery

The starting point is to extend the definition of type—think of the types of Definition 3.1 as extensional; we now add intensional ones.

Definition 4.1. Extensional and intensional types are characterized as follows.

1. 0 is an *extensional type*.
2. If $T_1, \ldots, T_n$ are types, extensional or intensional, $\langle T_1, \ldots, T_n \rangle$ is an *extensional type*.
3. If T is an extensional type, $\uparrow T$ is an *intensional type*.

Suppose, for example, that people are objects of type T. Then *Russian-population* would designate a property of type $\langle T \rangle$, with the intended value being the set of those people who are Russian citizens, and *Russian* would designate a property of type $\uparrow\langle T \rangle$, the intensional object associating different sets of people with different time points.

A higher-type modal language must be defined. I'll assume there is a distinct set of variables corresponding to each type, and variables are classified as intensional or extensional, according to their types. An extension-of operator on terms, $\downarrow$, is needed, just as in **FOIL**. A modal operator $\Box$ is added (with $\Diamond$ defined). Definition 3.2, for higher-type classical terms, formulas, and abstracts, must be modified to take modal operators and intensional objects into account. I only give the additions and changes to the earlier definition, and do not repeat what is not changed.

Definition 4.2. For formula, term, and predicate abstract, modifying Definition 3.2.

1. If Φ is a formula and $x_1, \ldots, x_n$ is a sequence of distinct variables of types $T_1, \ldots, T_n$ respectively, then $\langle \lambda x_1, \ldots, x_n.\Phi \rangle$ is a term of the intensional type $\uparrow\langle T_1, \ldots, T_n \rangle$.
2. Predicate abstracts and variables are terms. If t is a term of intensional type $\uparrow T$ then $\downarrow t$ is a term of type T.

3. If t is a term of either type $\langle T_1, \ldots, T_n\rangle$ or type $\uparrow\langle T_1, \ldots, T_n\rangle$, and t_1, $\ldots$, t_n is a sequence of terms of types T_1, $\ldots$, T_n respectively, then $t(t_1, \ldots, t_n)$ is an atomic formula.

4. Formulas are built up using propositional connectives and quantifiers exactly as in the classical case.

5. If Φ is a formula, so is $\Box\Phi$.

Case 2 uses the extension-of operator, converting a term of intensional type to a term of extensional type. For instance, if t is the intensional *Russian* example, $\downarrow t$ at a particular time instant should be the set of people who are Russian citizens at that moment. The symbols $\uparrow$ and $\downarrow$ were chosen to suggest that one 'cancels' the other. But note, $\downarrow$ is a symbol of the formal language, while $\uparrow$ is in the metalanguage—part of the typing mechanism.

Next, semantics. We need a modal analog of higher-type domains—Definition 3.3.

Definition 4.3. Let $\mathcal{G}$ be a non-empty set (of possible worlds) and let $\mathcal{D}$ be a non-empty set (the ground-level domain). For each type T, a collection $[\![T, \mathcal{D}, \mathcal{G}]\!]$ is defined as follows.

1. $[\![0, \mathcal{D}, \mathcal{G}]\!] = \mathcal{D}$.

2. $[\![\langle T_1, \ldots, T_n\rangle, \mathcal{D}, \mathcal{G}]\!]$ is the collection of all subsets of $[\![T_1, \mathcal{D}, \mathcal{G}]\!] \times \cdots \times [\![T_n, \mathcal{D}, \mathcal{G}]\!]$.

3. $[\![\uparrow T, \mathcal{D}, \mathcal{G}]\!]$ is the set of all functions from $\mathcal{G}$ to $[\![T, \mathcal{D}, \mathcal{G}]\!]$.

A member of $[\![T, \mathcal{D}, \mathcal{G}]\!]$ is an *object of type T over D and G*. An object is intensional or extensional or according to whether its type is intensional or extensional.

A *valuation* in $\mathcal{D}$ and $\mathcal{G}$ is a mapping v that assigns to each variable of type T some object of type T over $\mathcal{D}$ and $\mathcal{G}$.

Now, finally, simultaneous characterizations of truth for formulas, and designation for terms.

Definition 4.4. Let $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{D}\rangle$ be a structure, with $\mathcal{G}$ non-empty (the set of possible worlds), $\mathcal{D}$ non-empty (the ground-level domain), and $\mathcal{R}$ a binary relation on $\mathcal{G}$ (accessibility). For each valuation v in $\mathcal{D}$ and $\mathcal{G}$, an associated mapping $\hat{v}$ on terms is defined, and simultaneously the truth-status of a formula with respect to worlds and valuations is defined.

1. If x is a variable then $\hat{v}(x)$ is the constant function mapping each world to $v(x)$.
2. If t is a term of type $\uparrow T$ then $\hat{v}(\downarrow t)$ is the function mapping each world Γ to $v(t)(\Gamma)$.
3. Suppose $\langle \lambda x_1, \ldots, x_n.\Phi\rangle$ is a predicate abstract of type $\uparrow\langle T_1, \ldots, T_n\rangle$. Let f be the function that assigns to each world Γ the set

   $$\{\langle w(x_1), \ldots, w(x_n)\rangle \mid w \text{ agrees with } v \text{ on all variables except possibly } x_1, \ldots, x_n, \text{ and } \mathcal{M}, \Gamma \Vdash_w \Phi\}.$$

   Then, $\hat{v}(\langle \lambda x_1, \ldots, x_n.\Phi\rangle)$ is the constant function that maps each world to f.
4. For an atomic formula $t(t_1, \ldots, t_n)$,
   (a) If t is of extensional type, $\mathcal{M}, \Gamma \Vdash_v t(t, \ldots, t_n)$ provided $\langle \hat{v}(t_1)(\Gamma), \ldots, \hat{v}(t_n)(\Gamma)\rangle \in \hat{v}(t)(\Gamma)$.
   (b) If t is of intensional type, $\mathcal{M}, \Gamma \Vdash_v t(t_1, \ldots, t_n)$ provided $\mathcal{M}, \Gamma \Vdash_v (\downarrow t)(t_1, \ldots, t_n)$ (Reducing things to the previous case).
5. $\mathcal{M}, \Gamma \Vdash_v \neg\Phi$ if it is not the case that $\mathcal{M}, \Gamma \Vdash_v \Phi$.
6. $\mathcal{M}, \Gamma \Vdash_v \Phi \wedge \Psi$ if $\mathcal{M}, \Gamma \Vdash_v \Phi$ and $\mathcal{M}, \Gamma \Vdash_v \Psi$.
7. $\mathcal{M}, \Gamma \Vdash_v (\forall x)\Phi$ if $\mathcal{M}, \Gamma \Vdash_w \Phi$ for every valuation w that agrees with v on all variables except possibly x.
8. $\mathcal{M}, \Gamma \Vdash_v \Box\Phi$ if $\mathcal{M}, \Delta \Vdash_v \Phi$ for all $\Delta \in \mathcal{G}$ such that $\Gamma\mathcal{R}\Delta$.

This completes specification of the formal syntactic and semantic machinery. Unfortunately it is not enough. Just as with the classical version presented in Section 3, the logic defined by these semantics cannot have a complete proof system. And just as classically, the solution is to introduce 'non-standard' models, analogs of the classical Henkin models. This is not the place to get into details, which are somewhat complex. I refer you to [Fit02], where a full development can be found. The system given here, in turn, is a mild elaboration of an intensional higher-order logic of Montague and Gallin, [Mon60, Mon68, Mon70, Gal75]. The elaboration consists of introducing separate intensional and extensional types. In the Montague/Gallin system all types are intensional, though extensional objects can be introduced as intensions that do not depend on circumstances—rigid intensions, in other words.

4.2. Examples

With respect to a structure $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{D} \rangle$, there is an obvious distinction between *non-rigidity* and *rigidity*. A member, f, of an intensional type may take on different values at different worlds, in which case we say it is non-rigid. On the other hand, f might well be a constant function, in which case we say it is rigid. In general, the distinction between non-rigidity and rigidity is more than can be captured within our formal language, but a kind of localized version of it can be. Let us say f is *rigid at* Γ if the value that f has at Γ is the same as the value it has at any world accessible from Γ. This is something that can be said within our modal language, and I'll give two ways of doing so.

Let us suppose that, in addition to the machinery outlined in the previous section, we also have an equality predicate for each type, that is, there is a predicate $=$ of type $\langle T, T \rangle$, for each type T. (I know this violates what was said earlier about there being no constant symbols in the language, but please allow this extension, just this once.) I'll assume that $=$ is always interpreted as the equality relation, of appropriate type. The details are straightforward. Now, suppose $f \in [\![\uparrow T, \mathcal{D}, \mathcal{G}]\!]$, and v is a valuation mapping x to f. The following obviously says that f is locally rigid at Γ (in it, y is of type T).

$$\mathcal{M}, \Gamma \Vdash_v \langle \lambda y. \Box(y = \downarrow x) \rangle (\downarrow x)$$

There is another notion, essentially introduced by Gödel in the course of his ontological argument, which I have been calling *stability*. This time suppose g is not just intensional, but is in $[\![\uparrow\langle T \rangle, \mathcal{D}, \mathcal{R}]\!]$. If v maps z to g, the following is what it means to say g is *stable* at Γ (in it, w is of type T).

$$\mathcal{M}, \Gamma \Vdash_v (\forall w)[z(w) \supset \Box z(w)] \wedge (\forall w)[\Diamond z(w) \supset z(w)]$$

One can show that, for g in $[\![\uparrow\langle T \rangle, \mathcal{D}, \mathcal{R}]\!]$, stability at Γ and rigidity at Γ are equivalent. And this can be carried further still, to relate these notions to the vanishing of a *de re*, *de dicto* distinction.

In [Fit02], the higher-order intensional system sketched here was applied to analyze ontological arguments, Gödel's in particular. Subtle distinctions, such as whether Gödel intended intensional or extensional terms at various points in his argument, lead to seriously different results. If you are interested in seeing further examples of the present logic in use, I suggest looking at this material.

5. The Cumulative Set Hierarchy

At about the same time that Russell was developing type theory, Zermelo was introducing axiomatic set theory [Zer08, Zer35]. Behind the scenes, and serving as motivation for Zermelo, was a structuring of sets that has since been made explicit and has come to be known as the *cumulative hierarchy of sets*. This has long been a standard part of set theory courses. I'll just sketch the ideas here. Assume the underlying set theory is Zermelo-Fraenkel (Z-F).

One associates with each ordinal α a set, sometimes denoted V_α, sometimes R_α, as follows.

$$\begin{aligned} V_0 &= \emptyset \\ V_{\alpha+1} &= \text{power set of } V_\alpha \\ V_\lambda &= \cup_{\alpha<\lambda} V_\alpha \text{ for limit ordinals } \lambda \end{aligned}$$

This is not a separation into disjoint categories in the way that type theory is. It is cumulative, in the sense that $\alpha \leq \beta$ implies $V_\alpha \subseteq V_\beta$. Still, if x is in some V_α, there is a smallest ordinal β such that $x \in V_{\beta+1}$. In this case, x is said to be of *rank* β, and dividing the universe of sets up into groupings by rank amounts to sorting it out into things very much like types. If x is of rank β, the ranks of all members of x must be smaller ordinals. Rather nicely, the V_α themselves can be explicitly 'talked about,' since for each α, V_α is a member of $V_{\alpha+1}$; in fact, V_α itself has rank α. This allows the 'internalization' of things.

One can also define a limit—a proper class that is often denoted V_Ω—by setting $x \in V_\Omega$ if $x \in V_\alpha$ for some ordinal α. The members of V_Ω are called the *regular* or *well-founded* sets. It can be shown that if the Z-F axioms are consistent, they remain so when an axiom is added asserting that every set is regular. This additional axiom is quite commonly assumed, in fact. That is, it is standard to assume the universe of sets has a type-like structure, but one that is cumulative, which simplifies the construction of mathematical structures.

Whether mathematicians are explicit about their set-theoretic assumptions or not, it is a fact that essentially all of mathematics can be formalized within the framework sketched above. Indeed, standard mathematical objects tend to come along at 'low' levels in the cumulative hierarchy: the set of natural numbers (thought of as finite von Neumann ordinals) is in $V_{\omega+1}$, for instance. The cumulative hierarchy, and not type theory, has become the standard foundational structure for mathematics.

6. An Intensional Cumulative Set Hierarchy

A modal analog of the cumulative set hierarchy has been around for a long time, but has not generally been recognized to be that. Set-theoretic *forcing* was invented by Paul Cohen in the early 1960's, [Coh63, Coh64, Coh66], and gradually modified to its present form by a number of people, including Dana Scott and Robert Solovay, [SS71], and Joseph Shoenfield, [Sho71]. Although forcing was designed to be a powerful tool for set theorists, it can also be seen as providing us semantically with an intensional set theory. The presentation of forcing in [SF96] is in this direction, but the emphasis there is on the set-theoretic results, and not on the connections with intensional logic. Here I will say a bit more about these connections.

6.1. The General Case

The higher-order intensional system of types presented in Section 4 has both extensional and intensional types. This is convenient for the application for which it was designed, [Fit02]. But as noted earlier, it is also possible to formulate it with only intensional types—identifying extensional objects with rigid intensions. While one can take the either route with sets, it is standard to adopt an intension-only approach, and that is what will be done here.

Definition 6.1. Let $\mathcal{G}$ be a non-empty set, intended to be possible worlds. To each ordinal is associated a set, as follows.

$$\begin{aligned} V_0^{\mathcal{G}} &= \emptyset \\ V_{\alpha+1}^{\mathcal{G}} &= \text{ power set of } \mathcal{G} \times V_\alpha^{\mathcal{G}} \\ V_\lambda^{\mathcal{G}} &= \cup_{\alpha<\lambda} V_\alpha^{\mathcal{G}} \text{ for limit ordinals } \lambda \\ V_\Omega^{\mathcal{G}} &= \cup_\alpha V_\alpha^{\mathcal{G}} \text{ where the union is over all ordinals} \end{aligned}$$

Just as in the classical setting of Section 5, this gives us a cumulative hierarchy: $\alpha \le \beta$ implies $V_\alpha^{\mathcal{G}} \subseteq V_\beta^{\mathcal{G}}$, with a proper class, $V_\Omega^{\mathcal{G}}$ as limit. As before, we can define the rank of a member of $V_\Omega^{\mathcal{G}}$ to be the smallest ordinal α such that $f \in V_{\alpha+1}^{\mathcal{G}}$.

The idea behind the construction above is this: think of $f \in V_\Omega$ as encoding non-rigid membership information—if $\langle \Gamma, g\rangle \in f$, we should think of this as telling us that f has g as a member at world Γ. More formally, to $f \in V_{\alpha+1}^{\mathcal{G}}$ we associate a mapping, denoted f'', that assigns a subset of $V_\alpha^{\mathcal{G}}$ to each member of $\mathcal{G}$, by setting $f''(\Gamma) = \{g \in V_\alpha^{\mathcal{G}} \mid \langle \Gamma, g\rangle \in f\}$. Thus for each 'set' f, the associated mapping f'' is an *extension of* f function.

Assume we have a modal language with only ϵ as a binary relation symbol, written in infix position, and no constant or function symbols. (I use ϵ to distinguish intensional set membership from $\in$, the 'real' set membership relation.) There is only a single kind of quantifier. I'll set up modal models whose domain is the proper class $V_\Omega^\mathcal{G}$.

DEFINITION 6.2. Let $\mathcal{G}$ be a non-empty set, and $\mathcal{R}$ be a binary relation on $\mathcal{G}$, giving us a modal frame $\mathcal{M} = \langle \mathcal{G}, \mathcal{R} \rangle$. A *(set) valuation* in $\mathcal{M}$ is a mapping v from variables to members of $V_\Omega^\mathcal{G}$. Truth over $\mathcal{M}$ with respect to valuation v is defined as follows.

1. For atomic formulas, $\mathcal{M}, \Gamma \Vdash_v (x \,\epsilon\, y)$ iff $v(x) \in [v(y)]''(\Gamma)$.
2. The non-atomic cases are treated as usual in constant domain modal models, Definition 2.1, with quantifiers ranging over $V_\Omega^\mathcal{G}$.

One can obviously have $(x \,\epsilon\, y)$ true at one world but false at another, with respect to a valuation. That is, these are *intensional* sets. Among these intensional sets there will be some that are rigid. Indeed, we can identify those that are hereditarily rigid with the classical sets of V_Ω. Suppose we define a mapping from V_Ω to $V_\Omega^\mathcal{G}$ (making use of the well-foundedness of V_Ω) as follows. For each $d \in V_\Omega$:

$$\hat{d} = \mathcal{G} \times \{\hat{e} \mid e \in d\}$$

It is not hard to see that for each ordinal α, $d \in V_\alpha$ implies $\hat{d} \in V_\alpha^\mathcal{G}$. Also, $\hat{d}$ will be rigid, as will its members, members of members, and so on. And obviously, for $d, e \in V_\Omega$, we have $e \in d$ if and only if $\mathcal{M}, \Gamma \Vdash_v (x \,\epsilon\, y)$ when $v(x) = \hat{e}$ and $v(x) = \hat{d}$, independently of Γ.

Still, things can be rather badly behaved. One of the basic principles of classical set theory is that sets are completely determined by their membership—the so-called *axiom of extensionality.* It should not be a great surprise to find that this fails badly for intensional sets. Here is an example. It is standard in Z-F to identify numbers with certain sets—von Neumann ordinals—so that $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, and so on. Then there are rigid members $\hat{0}$, $\hat{1}$, and $\hat{2}$ in $V_\Omega^\mathcal{G}$, for every choice of $\mathcal{G}$. (In fact, these particular ones are all in $V_3^\mathcal{G}$.) Now, say $\mathcal{G}$ consists of just two worlds, Γ and Δ, and $\mathcal{M} = \langle \mathcal{G}, \mathcal{R} \rangle$, where the specification of $\mathcal{R}$ is not relevant. Consider the non-rigid set $f \in V_3^\mathcal{G}$ such that $f = \{\langle \Gamma, \hat{0} \rangle, \langle \Gamma, \hat{1} \rangle, \langle \Delta, \hat{1} \rangle\}$. Suppose v is a valuation such that $v(x) = \hat{2}$, $v(y) = f$, and $v(z) = \hat{3}$. Then we have $\mathcal{M}, \Gamma \Vdash_v (\forall w)[(w \,\epsilon\, x) \equiv (w \,\epsilon\, y)]$, that is, at Γ, $\hat{2}$ and f have the same members. But $\mathcal{M}, \Gamma \Vdash_v (x \,\epsilon\, z)$ while $\mathcal{M}, \Gamma \nVdash_v (y \,\epsilon\, z)$, so even though Γ

'thinks' $\hat{2}$ and f have the same members, it does not 'think' they are equal. Indeed, we should expect something like this because members of $\hat{3}$ must be rigid, but f is not.

The intensional set models constructed above have not been investigated in general. I do think a wider exploration would be of considerable interest. But it is time to turn to their most profound application.

6.2. The Important Special Case

Suppose we think of possible worlds as incomplete states of information, approximations. And suppose we think of a move from a possible world to an accessible possible world as a move that increases information, that improves the approximation. Naturally, as we move from world to accessible world to accessible world, we want 'facts' to accumulate. This forces two conditions on us. The first condition is syntactic: we should confine our attention to 'facts' of the form $\Box X$, because only these say something about arbitrary accessible worlds. The second condition is semantic: we should work with models in which accessibility is transitive (so 'facts' accumulate) and reflexive (so the 'facts' of any world from which we start a sequence of approximations remain with us). That is, we should work with $S4$ models.

Under $S4$ circumstances, positive information, represented by formulas of the form $\Box X$, does accumulate as we move from world to accessible world. But what about negative information—formulas of the form $\neg\Box X$? We cannot expect accumulation for them, of course, or the modal operator would become trivial. But suppose we weaken the requirement by asking only that, for negative information, there should at least be the possibility of it becoming permanent. This can be accomplished by further restricting formulas to the form $\Box\Diamond X$. Since these begin with a necessity symbol, we retain the accumulation of positive information. But also, since $\neg\Box\Diamond X \equiv \Diamond\Box\neg X$, even for negative information we do have the possibility of it becoming permanent—we have the possibility of $\Box\neg X$. All this is quite loose, but it suggests we consider the following mapping from a classical language (without modal operators) into a modal language.

Definition 6.3. For each formula X without modal operators, define a modal formula $\langle\!\langle X \rangle\!\rangle$ as follows.

1. $\langle\!\langle A \rangle\!\rangle = \Box\Diamond A$ for A atomic
2. $\langle\!\langle \neg X \rangle\!\rangle = \Box\Diamond\neg\langle\!\langle X \rangle\!\rangle$
3. $\langle\!\langle X \wedge Y \rangle\!\rangle = \Box\Diamond(\langle\!\langle X \rangle\!\rangle \wedge \langle\!\langle Y \rangle\!\rangle)$

4. $\langle\!\langle(\forall x)\Phi\rangle\!\rangle = \Box\Diamond(\forall x)\langle\!\langle\Phi\rangle\!\rangle$

This translation has a number of nice features. First of all, if other connectives and quantifiers are defined in the usual ways, they behave decently. For example, $\langle\!\langle X \supset Y\rangle\!\rangle \equiv \Box\Diamond(\langle\!\langle X\rangle\!\rangle \supset \langle\!\langle Y\rangle\!\rangle)$ is $S4$ valid, as is $\langle\!\langle(\exists x)\Phi\rangle\!\rangle \equiv \Box\Diamond(\exists x)\langle\!\langle\Phi\rangle\!\rangle$. Next, formulas are easily seen to be preserved under passage to alternative worlds, since $\langle\!\langle X\rangle\!\rangle \equiv \Box\langle\!\langle X\rangle\!\rangle$ is $S4$ valid. And negations behave as we wanted, since $\neg\langle\!\langle X\rangle\!\rangle \supset \Diamond\langle\!\langle\neg X\rangle\!\rangle$ is $S4$ valid—negative information has the possibility of becoming permanent. What is more, it can be shown that this is an exact embedding of classical logic into $S4$—X is classically valid iff $\langle\!\langle X\rangle\!\rangle$ is $S4$ valid, [Fit70, SF96].

So, from here on let us require accessibility relations to be reflexive and transitive—frames are for $S4$—and let us consider the behavior of formulas of the form $\langle\!\langle X\rangle\!\rangle$, where X is a formula in the language of classical Z-F set theory. Unfortunately, there is still the problem of the failure of the extensionality axiom—it is a problem that does not get solved this easily. Of course for many purposes one might not want an intensional version of set theory to obey an extensionality principle, but still it is pertinent to ask, what must be done to have it if we want it?

The ingenious solution to restoring the extensionality axiom for rigid and non-rigid sets alike is due to Paul Cohen, [Coh63, Coh64, Coh66], though his work was in a different context. The idea begins with the introduction of a carefully crafted candidate for an equality relation, leading in turn to a better candidate for the membership relation. Let $\approx$ be a new binary relation symbol—we want it to have the following behavior: for every $\Gamma \in \mathcal{G}$,

$$\begin{aligned}\mathcal{M}, \Gamma \Vdash_v (f \approx g) \equiv \langle\!\langle &(\forall x)[(x \epsilon f) \supset (\exists y)(y \epsilon g \wedge x \approx y)] \wedge \\ &(\forall x)[(x \epsilon g) \supset (\exists y)(y \epsilon f \wedge x \approx y)]\rangle\!\rangle\end{aligned}$$

It is not at all obvious that this can be achieved, but in fact it can—one introduces a series of approximations to it, $\approx_\alpha$ for each ordinal α, where $\approx_\alpha$ is defined in terms of $\approx_\beta$ with $\beta < \alpha$, and then one defines $\approx$ as a limit of these $\approx_\alpha$. Once one has this, one introduces a new version of membership, denoted here by $\hat{\in}$, by setting

$$\mathcal{M}, \Gamma \Vdash_v (f \mathrel{\hat{\in}} g) \text{ if } \mathcal{M}, \Gamma \Vdash_v (\exists x)\langle\!\langle x \approx f \wedge x \epsilon g\rangle\!\rangle$$

From now on, apply Definition 6.3 with the understanding that the classical atomic formula $(x \in y)$ is to be translated as $\Box\Diamond(x \mathrel{\hat{\in}} y)$. Introducing this more elaborate version of membership does, in fact, give us the validity of $\langle\langle X \rangle\rangle$, where X is the standard axiom of extensionality, and it does so even though we still have non-rigid intensional sets around.

What is truly remarkable, and in fact is the real significance of this construction, is that we get much more than just the axiom of extensionality. If $\mathcal{M} = \langle \mathcal{G}, \mathcal{R} \rangle$ is any $S4$ frame, and if X is any theorem of classical Z-F set theory, then $\langle\langle X \rangle\rangle$ is valid in $\mathcal{M}$. Still further, there are particular choices of $S4$ frames that invalidate $\langle\langle Y \rangle\rangle$, where Y is the continuum hypothesis. This means we have the classical unprovability of the continuum hypothesis! One way of looking at set-theoretic forcing is that it is the exploration of intensional models of set theory.

7. Conclusion

I have presented intensional versions of both higher-order classical logic, and set theory. Both ultimately date from the 1960's and, while intensional higher-order logic has been explored and utilized by a few logicians, not even that can be said for intensional set theory. It has been entirely the provence of classical set-theoreticians who think of it as something else altogether. It is time for these constructs to become better known, and in particular, for intensional set theoretic models to be explored for their own sakes.

References

[Coh63] Paul J. Cohen. The independence of the continuum hypothesis, I. *Proceedings of the National Academy of Sciences, U.S.A.*, 50:1143–1148, 1963.

[Coh64] Paul J. Cohen. The independence of the continuum hypothesis, II. *Proceedings of the National Academy of Sciences, U.S.A.*, 51:105–110, 1964.

[Coh66] Paul J. Cohen. *Set Theory and the Continuum Hypothesis*. W. A. Benjamin, New York, 1966.

[Fit70] Melvin C. Fitting. An embedding of classical logic in S4. *Journal of Symbolic Logic*, 35:529–534, 1970.

[Fit01] Melvin C. Fitting. First-order intensional logic. Forthcoming in *Proceedings of the Tarski Conference*, 2001.

[Fit02] Melvin C. Fitting. *Types, Tableaus, and Gödel's God*. Kluwer, 2002.

[FM98] Melvin C. Fitting and Richard Mendelsohn. *First-Order Modal Logic*. Kluwer, 1998. Paperback, 1999.

[Gal75] D. Gallin. *Intensional and Higher-Order Modal Logic*. North-Holland, 1975.

[HC96] G. E. Hughes and M. J. Cresswell. *A New Introduction to Modal Logic*. Routledge, London, 1996.

[KK01] Marcus Kracht and Oliver Kutz. The semantics of modal predicate logic I, counterpart frames. In Frank Wolter, Heinrich Wansing, Maarten de Rijke, and Michael Zakharyaschev, editors, *Advances in Modal Logic*, volume 3, Stanford, 2001. CSLI Publications.

[KK03] Marcus Kracht and Oliver Kutz. The semantics of modal predicate logic II, modal individuals revisited. In Reinhard Kahle, editor, *Intensionality. Proceedings of a conference held at Munich, 27th–29th October 2000*, Lecture Notes in Logic. A. K. Peters, 2003. To appear.

[Kra93] Marcus Kracht. How completeness and correspondence theory got married. In M. de Rijke, editor, *Diamonds and Defaults*, Dordrecht, 1993. Kluwer.

[Kut02] Oliver Kutz. New semantics for modal predicate logics. In *Foundations of the Formal Sciences II*, Trends in Logic. Kluwer, 2002.

[Lew68] David Lewis. Counterpart theory and quantified modal logic. *Journal of Philosophy*, 65:113–126, 1968.

[Lew71] David Lewis. Counterparts of persons and their bodies. *Journal of Philosophy*, 68:203–211, 1971.

[Mon60] Richard Montague. On the nature of certain philosophical entities. *The Monist*, 53:159–194, 1960. Reprinted in [Tho74], 148–187.

[Mon68] Richard Montague. Pragmatics. pages 102–122. 1968. In *Contemporary Philosophy: A Survey*, R. Klibansky editor, Florence, La Nuova Italia Editrice, 1968. Reprinted in [Tho74], 95–118.

[Mon70] Richard Montague. Pragmatics and intensional logic. *Synthèse*, 22:68–94, 1970. Reprinted in [Tho74], 119–147.

[Rus05] Bertrand Russell. On denoting. *Mind*, 14 (new series):479–493, 1905.

[Rus08] Bertrand Russell. Mathematical logic as based on the theory of types. *American journal of mathematics*, 30:222–262, 1908. Reprinted in [vH67], pp 150-182.

[Sco71] Dana S. Scott, editor. *Axiomatic Set Theory I*, volume 13 of *Symposium in Pure Mathematics*, Providence, R.I., 1971. American Mathematical Society. (Held in Los Angeles, 1967).

[SF96] Raymond M. Smullyan and Melvin C. Fitting. *Set Theory and the Continuum Problem*. Oxford University Press, 1996.

[Sho71] Joseph R. Shoenfield. Unramified forcing. pages 357–381, 1971. In [Sco71].

[SS71] Dana S. Scott and Robert M. Solovay. Boolean-valued models for set theory. 1971. In [Sco71].

[Tho74] Richmond H. Thomason, editor. *Formal Philosophy, Selected Papers of Richard Montague*. Yale University Press, New Haven and London, 1974.

[vH67] Jean van Heijenoort, editor. *From Frege to Gödel*. Harvard University Press, Cambridge, MA, 1967.

[WR27] Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*. Cambridge University Press, Cambridge, 2nd edition, 1927. First edition, 1910.

[Zer08] Ernst Zermelo. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, 65:261–281, 1908. English translation in [vH67], pp 199-215.

[Zer35] Ernst Zermelo. Grundlagen einer allgemeinen Theorie der mathematischen Satzsysteme. *Fundamenta Mathematicae*, 25:136–146, 1935.