First-order Modal Logic

Pablo Rivas-Robledo, Magnus Kjærgaard, Ludovico Ageno

Contents

1 Motivation: Why Quantify Modal Logic?* 2
1.1 Opaque Contexts: Modalities *de dicto* and *de re* . . . 2
1.2 The Distinction Itself . . . 4
1.3 Proper Names, Definite Descriptions, and Rigid Designation . . . 5
1.4 The Shape of the Problem to Come . . . 7

2 The Naïve Approach 9
2.1 Syntax* . . . 9
2.2 Semantics . . . 10
2.3 Working in the Naïve System . . . 12
2.3.1 A worked example . . . 12
2.3.2 The Barcan formulae and their failure . . . 13
2.3.3 Identity and the necessity of identity . . . 16
2.4 The Failure of Completeness, Intuitively † . . . 17

3 The First Response: Necessitism 22
3.1 The Fix: Hold the Domain Fixed . . . 23
3.2 A Refinement: Kripke Models with Equality . . . 24
3.3 Reclaiming Completeness† . . . 24
3.4 The Commitment: Necessitism . . . 25

4 The Second Response: Contingentism 29
4.1 The Semantics: Varying Domains and Free Logic . . . 29
4.2 The Cost: Universal Instantiation Fails . . . 31
4.3 The First Payoff: Both Barcan Formulae Fail . . . 32
4.4 The Second Payoff: Completeness Recovered† . . . 33
4.5 A Variation and a Scruple . . . 34

Remark: Sections marked with a * are background and sections marked with a † are optional/advanced material.

Acknowledgements: This is supplementary material for Chapter 3 of L. T. F. Gamut's *Logic, language, and meaning. Volume II. Intensional logic and logical grammar* but can be read as a stand-alone text provided one is familiar with first-order logic and relational semantics for modal logic.

1 Motivation: Why Quantify Modal Logic?*

In the propositional modal logic of the previous chapters we learned to reason with the operators $\Box$ and $\Diamond$, but the sentences over which they ranged were structureless: p, q, $\Box p$, $\Diamond(p \to q)$, and the like. Natural language, of course, is not structureless. Its sentences contain names that pick out individuals, predicates that ascribe properties and relations, and quantifiers that talk about *some* and *all*. The moment we wish to say not merely *that something is possible* but *which thing is possibly which way*, propositional modal logic falls silent. We need the resources of predicate logic and the resources of modal logic at once. The system that results from adding modal (or, more generally, intensional) operators to predicate logic is *quantified intensional logic*, or, in the special case on which we shall concentrate, *first-order modal logic*.

The merger looks innocent. One simply writes $\Box$ and $\Diamond$ in front of formulas of predicate logic and lets quantifiers and operators stand in whatever configuration the grammar of a sentence demands. But the innocence is deceptive. The interaction between quantifiers, which talk about individuals, and modal operators, which talk about possibilities, raises a cluster of questions that have no analogue in either parent system. *Do the same individuals exist in every possible world? Does a name refer to the same thing from world to world? What does it mean for an individual in one world to be the same as an individual in another?* These questions reach into some of the most contested terrain in the philosophy of language and metaphysics. Before we begin building the semantics in earnest (a project that will occupy the sections to follow) this section lays out what is at stake. We motivate the enterprise by displaying the linguistic phenomena that call for it, and we survey the philosophical debates into which any adequate treatment of those phenomena will inevitably lead.

1.1 Opaque Contexts: Modalities *de dicto* and *de re*

As a point of entry, we return to a theme from §1.6, where intensional constructions were introduced as constructions for which certain substitution principles of predicate logic fail. Following Quine (1953), we say that intensional constructions create *opaque* contexts. These are to be distinguished from *transparent* contexts, for which the substitution principles hold. The principle at issue is the one licensing substitution of identicals, schematically (1):

$$s = t \models \varphi \leftrightarrow [t/s]\varphi \tag{1}$$

In a transparent context, if s and t refer to the same thing then replacing s by t in φ leaves the truth value of φ untouched. In an opaque context it need not. Here are some constructions giving rise to opaque contexts, each accompanied by an example showing that (1) is violated.

Quotation

Sentence (3) does not follow from (1) and (2):

(1) The sailor spoke the words *Ahoy, Captain*.

(2) The Captain is Hipólito.

(3) The sailor spoke the words *Ahoy, Hipólito*.

Indirect speech

Sentence (6) does not follow from (4) and (5):

(4) Amina said that Kwame solved the puzzle.

(5) Kwame is the youngest student in the class.

(6) Amina said that the youngest student in the class solved the puzzle.

Propositional attitudes

Constructions with verbs like *to discover*, *to believe*, *to suspect*, and *to know*. Sentence (9) does not follow from (7) and (8):

(7) Inspector Mehta knows that the forger signed the ledger.

(8) Rosa is the forger.

(9) Inspector Mehta knows that Rosa signed the ledger.

Verbs of intention

Constructions with verbs such as *to look for* and *to wish for*. The example from §1.6: sentence (12) does not follow from (10) and (11):

(10) John is looking for the supreme commander of the armed forces of the United States of America.

(11) The president of the United States of America is the supreme commander of the United States armed forces.

(12) John is looking for the president of the United States of America.

Modality

It is a necessary truth that nine exceeds seven: given the meanings of *nine*, *seven*, and *exceeds*, the sentence *Nine exceeds seven* could not have been false. But surely (13) is no necessary truth:

(13) The number of planets exceeds seven.

The truth or falsity of (13) is not settled by the meanings of the expressions it contains. It expresses a contingent astronomical fact; that there are more than seven planets is something discovered by observation and inference. So (16) does not follow from (14) and (15):

(14) Nine necessarily exceeds seven.

(15) Nine is the number of planets.

(16) The number of planets necessarily exceeds seven.

Besides these, there are many more constructions giving rise to opaque contexts. However, we will only focus the modal case, for modal predicate logic is the oldest and best-understood member of the intensional family, and what we learn about it transfers, *mutatis mutandis*, to its temporal, epistemic, and deontic relatives.

1.2 The Distinction Itself

Philosophers have reacted in different ways to the failure of substitution in opaque contexts. Let us dwell on (16). It might be argued that there is a reading of (16) on which it *does* follow from (14) and (15). This reading may be paraphrased: *that number which is in fact the number of planets is necessarily greater than seven.* It translates as (17), whereas the reading on which (16) does *not* follow may be rendered as (18):

(17) $\exists x(x = \text{the number of planets} \wedge \Box(x > 7))$

(18) $\Box\exists x(x = \text{the number of planets} \wedge x > 7)$

Reading (18) says that in every possible situation the number of planets, whatever it happens to be, exceeds seven. These two readings lead to a distinction traditionally drawn in modal logic, between modalities *de dicto* and *de re*. In a predicate language with an added $\Box$ this distinction can be expressed in terms of the *scope* of $\Box$. Consider the simpler examples (19) and (20) and their translations (21) and (22):

(19) Necessarily there is something greater than seven.

(20) There is something necessarily greater than seven.

(21) $\Box\exists x(x > 7)$

(22) $\exists x\,\Box(x > 7)$

In (21) the scope of $\Box$ is $\exists x(x > 7)$; in (22) it is $x > 7$. If every variable within the scope of $\Box$ is bound by a quantifier likewise within its scope, then $\Box$ is a modality *de dicto*: examples are (18) and (21). If, on the other hand, a variable within the scope of $\Box$ is bound by a quantifier *outside* that scope so that the variable occurs free in the operator's scope, then $\Box$ is a modality *de re*: examples are (17) and (22).

The formal distinction matches a traditional one. A modality *de dicto* was seen as the attribution of necessary (or possible) truth to a proposition (a *dictum*); a modality *de re* as the attribution of a necessary (or possible) property to a thing (a *res*). In asserting (21), one asserts that the proposition $\exists x(x > 7)$ is necessarily true; in asserting (22), one asserts the existence of a thing that necessarily *has* the property of being greater than seven.

Some philosophers have balked at the second kind. To recognize modalities *de re*, they hold, is to revive *essentialism*: the doctrine that distinguishes a thing's accidental properties from its essential ones. Having objections to essentialism, they reject *de re* modality as meaningless, or at best urge that it be reduced to *de dicto* modality. The most vigorous such opponent was Quine. Even setting aside whether *de re* modality really does entail essentialism, a position like Quine's seems to us ill-suited to our purposes. As we shall insist more than once, philosophical objections should not weigh heavily when the aim is the description of natural language. We want descriptions of how we *do* speak, not of how we would have to speak to earn a philosopher's approval. Speakers may well make philosophically dubious assumptions; that is a fact about them, not a defect to be reformed away. And that *de re* modalities occur in natural language seems to us indisputable. Consider (23):

(23) Each of those present may have committed the murder.

What this means is clear. It may be formalized $\forall x\,\Diamond Mx$, and it plainly does *not* mean the same as $\Diamond\forall x\,Mx$, which formalizes (24):

(24) It is possible that each of those present committed the murder.

It is far from clear how the *de re* reading (23) could be reduced to the *de dicto* reading (24). As we shall see, possible-worlds semantics supplies a perfectly clear interpretation for *de re* modality, so there is no need to do without it.

The same ambiguities arise once we add tense operators, like **P** to make statements about the past. Sentence (25) has two readings, (26) and (27), with the key in (28):

(25) The president received a decoration.

(26) $\mathbf{P}(\exists x(\forall y(Py \leftrightarrow y = x) \wedge Dx))$

(27) $\exists x(\forall y(Py \leftrightarrow y = x) \wedge \mathbf{P}(Dx))$

(28) Px: x is president; Dx: x receives a decoration.

Here too (26) is the *de dicto* reading and (27) the *de re* reading. The same distinction, the same role for scope.

1.3 Proper Names, Definite Descriptions, and Rigid Designation

The *de dicto*/*de re* distinction shows that quantified modal logic is needed to render the readings natural-language sentences actually have. But it also opens onto a second and deeper motivation once we ask how a singular term behaves across possible worlds. We therefore step directly into one of the central debates of twentieth-century philosophy of language: the debate over proper names and definite descriptions.

A definite description, such as *the number of planets* or *the president of the United States*, can clearly pick out different things in different circumstances. That is exactly why (16) had its two readings: were the reference of *the number of planets* the same in every world, it would everywhere be nine, and the *de dicto* reading (18) would come out true, contrary to what we concluded. Indeed the entire *de dicto*/*de re* distinction would collapse for a term whose reference never varied. So at least *some* singular terms must be allowed to shift their reference from world to world.

The pressing question is whether *proper names* are like this too. Frege, as we saw in §1.7, held that every name, proper names included, has a *Sinn* expressible by a definite description. To know the meaning of *Aristotle* is to associate with it some description: *the teacher of Alexander*, say, or *the author of the* Metaphysics. On this *description theory* of names, a proper name is shorthand for a description, and inherits the description's capacity to shift its reference across worlds.

Description theory explains how a name manages to refer to one individual rather than another: the associated description supplies properties that single that individual out. It explains why *Hesperus is Hesperus* and *Hesperus is Phosphorus* differ in meaning, even though both names refer to Venus, for the two names go proxy for *different* descriptions. And it makes room for true negative existentials. *Pegasus does not exist* looks paradoxical if *Pegasus* is a name that must refer to something in order to be meaningful. But if *Pegasus* abbreviates *the winged horse*, then the sentence says merely that there is nothing in the world that is a winged-horse, which can be true without commitment to any nonexistent steed.

Yet, the description theory has well-known costs, and it is in articulating those costs that the alternative picture takes shape. Different speakers associate different descriptions with the same name, and some associate mistaken ones; so the theory threatens to make

the meaning of a name vary from speaker to speaker, and to make competent users *wrong* about whom they are talking about.

Worse, the descriptions we associate with a name seem not to be true of its bearer *necessarily*. Suppose everything we believe about Laozi (that he served as an archivist in the Zhou court, that he composed the *Dào Dé Jīng* in a single sitting before departing to the western wilderness) turned out to misdescribe the actual person who bore the name. We would not conclude that Laozi never existed; we would conclude that we had been *wrong about Laozi*. The name seems to keep hold of its bearer even as the descriptions fall away. Attempts to repair the theory by taking names to abbreviate *disjunctions* or *clusters* of descriptions (Searle, 1958) run into the awkward question of how many is enough, and whether some descriptions should count for more than others. The repairs raise more problems than they settle.

These difficulties motivated Kripke's (1980) influential alternative. Kripke held that there is a fundamental difference between the two kinds of expressions. Definite descriptions may change their reference from world to world; proper names do not. A proper name refers to the *same* thing in every world in which it refers at all. Such a term Kripke called a *rigid designator*. We illustrated the difference in §1.7 with a counterfactual:

(29) If Nakamura had won last year's championship, then the world champion would have been an amateur.

In the worlds the antecedent invites us to consider, *Nakamura* refers to the very person it refers to here, while *the world champion* refers to someone other than its actual-world referent. The name is rigid; the description is not.

Kripke's proposal has consequences that will run through the sections to come. The first concerns *transworld identity*. Possible-worlds semantics seems to raise an complicated question: what is it for an individual in one world to be *the same* individual as one in another? If we had to settle this by comparing properties, we would be in trouble, for an individual's properties may differ wildly from world to world. Kripke's answer is that the question is a pseudo-problem. We do not *discover* transworld identities by inspection; we *stipulate* them. In framing (29) we simply introduce worlds in which Nakamura, that person, wins, however much they may otherwise differ from the actual Nakamura. The name carries the identity with it; no comparison of properties is required or even intelligible.

The second consequence concerns identity statements themselves. If names are rigid, then a true identity between names, such as *Hesperus is Phosphorus*, would seem to be *necessarily* true: both names rigidly designate Venus, so they co-designate in every world. Kripke embraced this. The necessity of an identity, he urged, has nothing to do with how we come to *know* it. He drew apart two notions long run together: the *a priori*, an epistemological notion (a truth knowable independently of experience), and the *necessary*, a metaphysical one (a truth that could not have been otherwise). *Hesperus is Phosphorus* is necessary but not a priori, its truth was an astronomical discovery. *Hesperus is Hesperus* is both. So the necessity of identity does not collapse the hard-won distinction Frege drew between the two. This thesis (the necessity of identity) is one we will be able to state precisely and assess *within* the formal semantics in Section 3.

The third consequence concerns existence. A name like *Pegasus* can fail to denote anything that exists, yet we want *Pegasus does not exist* to come out true and, on Kripke's view, we cannot reach for the description-theoretic paraphrase to get it. How a rigid-

designation semantics is to accommodate non-existence, and what an *existence predicate* might do for us, is a question we postpone to Section 4.

We do not mean to settle the names debate here; that is not our task. What matters for us is that the choice between description and rigid designation is not merely a philosophical one. It is a choice about how the formal semantics of our singular terms should be set up: whether constants are to be interpreted as *individuals*, fixed across worlds, or as *individual concepts*, functions from worlds to individuals that may vary. Each option carries commitments, and the commitments will shape everything that follows.

1.4 The Shape of the Problem to Come

We have now seen two families of motivation for quantifying modal logic. The first is descriptive: natural-language sentences exhibit *de dicto*/*de re* ambiguities that only the interplay of quantifiers and modal operators can capture. The second is foundational: the moment we let singular terms loose in modal contexts, we confront the debates over names, descriptions, rigid designation, transworld identity, and existence. A semantics for quantified modal logic is, in effect, a set of answers to these debates.

This is the central lesson of the sections that follow, and it is worth stating plainly in advance. There is no single, neutral, obviously correct semantics for quantified modal logic. The naïve approach (simply placing modal operators over ordinary predicate logic, and reading off truth conditions in the obvious way) turns out to be untenable, as we show in Section 2. The reasons it fails are instructive, for they reveal that we are forced to make choices: about whether the domains of different worlds may differ, about how names refer across worlds, about what the quantifiers range over, about whether everything that exists exists *necessarily*. And these choices do not float free of one another. They lock together into coherent positions, each with a price.

Two such positions will concern us. On the first, developed in Section 3, we keep a classical predicate logic, add an identity predicate, and pay for the convenience with *necessitism*: the thesis that necessarily everything necessarily exists. On the second, developed in Section 4, we move instead to a *free* logic, in which terms may fail to denote and quantifiers do not presuppose existence, and we thereby earn the right to *contingentism*: the thesis that what exists is a contingent matter, that there might have been other or fewer things. Each position keeps something we want and gives up something we want; neither is cost-free. The task of the remaining sections is to set out the two packages with enough precision that the reader can see exactly what each costs and exactly what each buys.

Throughout, our guiding commitment remains the one announced above. We are describing natural language, not legislating metaphysics. Where a philosophical objection would bar a construction that speakers plainly use, we set the objection aside rather than reforming the speakers. But we will also find that the formal choices themselves *carry* metaphysical commitments, whether we like it or not, and this is the interest of the necessitism/contingentism division. One cannot build a working semantics for quantified modal logic without taking a stand. The most we can do, and what we shall try to do, is to make every stand explicit, and to count its cost honestly.

Exercises

Exercise 1

Translate the following sentences into formulas of intensional predicate logic.

(a) Osaka may win the Australian Open one day.

(b) Perhaps everybody has always been aware of something.

(c) Perhaps there is something that everybody has always been aware of.

(d) If anybody can be smarter than anybody else, then everybody can be the smartest.

(e) You can fool some of the people all of the time, and all of the people some of the time, but you can't fool all of the people all of the time.

(f) The president will always be a democrat. (Find two translations.)

(g) Every schoolchild believes that a mathematician wrote *Through the Looking Glass*. (Find three translations, representing 'x believes that φ' as $B(x, \varphi)$.)

Exercise 2

Find examples of intensional expressions (expressions that do not allow substitution of materially equivalent expressions in their scope) in each of the following categories:

(a) adjectives;

(b) adverbs;

(c) conjunctions;

(d) prepositions;

(e) determiners.

Exercise 3

For each of the sentences below, say whether the most natural reading is *de dicto* or *de re*, give a formalization, and explain how the scope of the modal or tense operator determines the reading.

(a) The richest person in the world might have been poor.

(b) Necessarily, the richest person in the world is rich.

(c) The next secretary-general will once have been a child.

Exercise 4*

Kripke distinguishes the *necessary* from the *a priori* and offers examples that fall under one notion but not the other. (a) Give, in your own words, a sentence that is plausibly necessary but not a priori, and explain why it has each status. (b) Goldbach's conjecture states that every even number greater than 2 is the sum of two primes. Assuming it is in fact true, explain why it would be necessary; and explain why, as things stand, we have no a priori knowledge of it. (c) What would have to change for us to acquire such knowledge?

2 The Naïve Approach

We have argued that natural language calls for a logic in which quantifiers and modal operators interact, and we have seen that any such logic must take a stand on a cluster of questions about how individuals behave across possible worlds. It is time to build one. In this section we build the most natural one there is, the one a newcomer would write down without a second thought, and we follow that thought to its consequences.

The guiding idea is this. A possible world is a way things could have been, and different ways things could have been may simply contain different things. The world might have held one more person than it does, or one fewer; there might have been objects that in fact never formed, and there might have failed to be objects that in fact exist. Things come into being and pass away, and what exists at one moment, or in one circumstance, need not exist in another. Nothing in the bare notion of a possible world rules any of this out, and so the most natural semantics lets each world carry its own domain of individuals, populated entirely as that world sees fit, with no constraint tying one world's population to another's. We take a propositional Kripke frame, with its worlds and its accessibility relation, and to each world we attach a domain of its own. Quantifiers at a world range over that world's domain; the modal operators travel between worlds as before. This is the naïve approach, and its naïveté lies exactly here, in the unreflective and genuinely appealing thought that worlds may be populated independently of one another.

The reader should bear in mind throughout that 'naïve' is not a term of abuse. The thought is a good one, arguably the correct starting point, and the semantics it yields is clean and workable. What we will discover is not that the approach is incoherent but that it carries consequences, that letting domains vary freely makes certain cherished principles fail, and that whether those failures are welcome or not is precisely what divides the positions we will study later. Our task in this section is to set the system up with care and become fluent in it. We give its syntax and its semantics, compute with it, and read off which principles it validates and which it does not. The critical scrutiny comes afterward.

2.1 Syntax*

This should just be a refresher. The language of quantified modal logic is obtained from the language of predicate logic by adding the modal operators.

Definition 1 (Alphabet)**.** *The alphabet of the language $\mathcal{L}_{\mathbf{FOML}}$ consists of:*

- *a countable set of individual variables, $x, y, z, \ldots$;*
- *a countable set of individual constants, $c_1, c_2, \ldots$;*
- *for each $n \geq 1$, a countable set of n-ary predicate letters, $P, Q, R, \ldots$;*
- *the connectives $\neg$ and $\rightarrow$;*
- *the quantifier $\forall$;*
- *the modal operator $\Box$;*
- *the parentheses (and).*

We call a variable or a constant a *term*, and we write $t, t_1, t_2, \ldots$ for arbitrary terms. The remaining connectives $\wedge$, $\vee$, $\leftrightarrow$, the existential quantifier $\exists$, and the possibility operator $\Diamond$ are introduced as abbreviations in the usual way: $\Diamond\phi$ abbreviates $\neg\Box\neg\phi$, and $\exists x\,\phi$ abbreviates $\neg\forall x\,\neg\phi$, just as in their respective parent logics.

Definition 2 (Formulas)**.** *The formulas of $\mathcal{L}_{\mathbf{FOML}}$ are given by the following BNF:*

$$\phi ::= P(t_1, \ldots, t_n) \mid \neg\phi \mid (\phi \to \psi) \mid \forall x\,\phi \mid \Box\phi,$$

where P is an n-ary predicate letter and $t_1, \ldots, t_n$ are terms.

The notions of *free* and *bound* occurrences of a variable, and of a *sentence* (a formula with no free variables), are exactly as in predicate logic; the modal operators bind nothing and so do not affect them. It is worth noticing already, however, that the scope of a $\Box$ may contain a variable that is free in that scope while being bound by a quantifier standing outside the $\Box$. This is the syntactic shape of a modality *de re*, in the sense of Section 1.2, and the whole interest of the semantics to come lies in how it treats such a configuration. In $\exists x\,\Box P(x)$ the variable x occurs free within the scope of $\Box$; in $\Box\exists x\,P(x)$ it does not. The two formulas are built from the same material in a different order, and we will find that the order makes all the difference.

2.2 Semantics

We come now to the part that matters. The question is what it should mean for a formula of $\mathcal{L}_{\mathbf{FOML}}$ to be true, and the answer follows the guiding idea without compromise: each world carries its own domain, and these domains are wholly independent of one another.

From propositional modal logic we take a *frame*, a nonempty set W of worlds with an accessibility relation R. To this we add a *system of domains*, one nonempty set D_w for each world w, the individuals that exist at that world. We impose *no* relation between the domains of distinct worlds. As we move from a world to one accessible from it, the population may grow, it may shrink, individuals may appear and others vanish, and all of this is permitted freely. This is what we mean by *varying domains*, and it is the most faithful rendering of the thought that different worlds may simply contain different things.

Definition 3 (Kripke Frame with Varying Domains)**.** *A* Kripke frame with varying domains *is a triple $\mathfrak{F} = \langle W, R, D\rangle$ where $\langle W, R\rangle$ is a Kripke frame and $D = \bigcup_{w\in W} D_w$ assigns to each world w a nonempty domain D_w, with no constraint relating D_w and D_v for distinct w, v.*

The predicates are interpreted world by world, as they must be if anything is to be contingent: an n-ary predicate P receives at each world w an extension $I(P, w) \subseteq (D_w)^n$, drawn from the individuals available there. For the constants we make the choice argued for in Section 1.3: a constant is a *rigid designator*, denoting one fixed individual, the same at every world.

Definition 4 (Model)**.** *A* model *is a tuple $\mathfrak{M} = \langle W, R, D, I\rangle$ where $\langle W, R, D\rangle$ is a frame with varying domains and I assigns*

- *to each constant c an individual $I(c) \in \bigcup_{w\in W} D_w$, the same at every world;*

- *to each n-ary predicate P and world w an extension $I(P, w) \subseteq (D_w)^n$.*

Letting domains vary forces a decision we did not have to make before. When we evaluate $\Box\phi$ at w we carry our assignment to an accessible world v, and a variable may there denote an object that does not belong to D_v. What, then, is the truth value of an atomic formula about an object that does not exist at the world of evaluation? We take the simplest answer available, and reserve any refinement for later. Assignments are allowed to draw on the union $\bigcup_w D_w$ of all the domains, so that a variable always has *some* value; and the extension $I(P, w)$ is a set of tuples from D_w only. An atomic formula $P(t_1, \ldots, t_n)$ is then *true* at w exactly when the tuple of values lies in $I(P, w)$, and *false* otherwise, including the case where some value does not belong to D_w at all. Truth remains total and entirely classical: nothing is left undefined, and a predication about an absent object simply comes out false. We will see in the section on contingentism that this crisp answer is open to challenge, and that taking the challenge seriously leads to free logic; but for the naïve approach it serves perfectly, and it keeps the logic classical while we explore the consequences of varying the domains.

Definition 5 (Assignment)**.** *An* assignment *is a function g mapping each variable into $\bigcup_{w \in W} D_w$. We write $g[x \mapsto d]$ for the assignment agreeing with g everywhere except that it sends x to d. The value of a term is $[\![c]\!]^{\mathfrak{M},g} = I(c)$ for a constant and $[\![x]\!]^{\mathfrak{M},g} = g(x)$ for a variable; the world plays no role in the value of a term.*

Definition 6 (Truth)**.** *Let $\mathfrak{M}$ be a model, w a world, and g an assignment. The relation $\mathfrak{M}, w \Vdash_g \phi$ is defined as follows:*

- *$\mathfrak{M}, w \Vdash_g P(t_1, \ldots, t_n)$ iff $\langle [\![t_1]\!]^{\mathfrak{M},g}, \ldots, [\![t_n]\!]^{\mathfrak{M},g} \rangle \in I(P, w)$;*
- *$\mathfrak{M}, w \Vdash_g \neg\phi$ iff $\mathfrak{M}, w \nVdash_g \phi$;*
- *$\mathfrak{M}, w \Vdash_g \phi \to \psi$ iff $\mathfrak{M}, w \nVdash_g \phi$ or $\mathfrak{M}, w \Vdash_g \psi$;*
- *$\mathfrak{M}, w \Vdash_g \Box\phi$ iff for every v with wRv, $\mathfrak{M}, v \Vdash_g \phi$;*
- *$\mathfrak{M}, w \Vdash_g \forall x\, \phi$ iff for every $d \in D_w$, $\mathfrak{M}, w \Vdash_{g[x \mapsto d]} \phi$.*

A formula is valid *if true at every world of every model under every assignment.*

Two clauses carry the weight of everything that follows, and the tension between them is the engine of the whole approach. The clause for $\Box$ moves us to an accessible world v *carrying the same assignment*, so a value d fixed for x at w remains the value of x at v, whether or not d still exists at v. But the clause for $\forall$ ranges only over D_w, the domain *of the world where the quantifier sits.* So a quantifier standing outside a box and a quantifier standing inside it are read against *different* domains, the outer one against the world of evaluation and the inner one against an accessible world that may have lost old individuals or gained new ones. In the naïve approach the box and the quantifier no longer agree about which individuals there are. That disagreement is the source of everything interesting in this section.

As in propositional modal logic, the truth definition does not by itself single out one logic. It says what it is for a formula to be true at a world, but it leaves the accessibility relation R unconstrained, and it is the constraints we place on R that decide which formulas come out

valid. If we impose nothing at all, ranging over the class of *all* frames with varying domains, we obtain the minimal quantified modal logic, the first-order counterpart of the propositional system **K**. We call it **QK**. Its validities are exactly the formulas true on every such frame: the predicate-logical validities, the distribution schema $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$ inherited from **K**, and what follows from these together with the quantifier clauses.

Stronger systems come, just as in the propositional case, from constraining R, and each constraint answers to a familiar axiom. Requiring R to be reflexive validates $\Box\phi \to \phi$; requiring it transitive validates $\Box\phi \to \Box\Box\phi$; requiring it symmetric validates $\phi \to \Box\Diamond\phi$. Imposing reflexivity and transitivity together gives the first-order counterpart of **S4**, which we call **QS4**; adding symmetry on top, so that R is an equivalence relation, gives **QS5**.[1] In general, for a propositional modal logic $\mathbf{\Lambda}$ we write $\mathbf{Q\Lambda}$ for its first-order counterpart, the system whose validities are the formulas true on the frames $\mathbf{\Lambda}$ describes.

Of all these it is **QS5** that will matter most to us in the end. An equivalence relation where every world sees each other, and this is exactly the picture one wants if $\Box$ is to mean truth in *every* possibility without restriction, which is the reading appropriate to metaphysical necessity. So **QS5** is the system on which we will focus.

2.3 Working in the Naïve System

With the semantics in hand we put the system to work, computing truth values and building small countermodels by hand, in the manner that will become our habit.

2.3.1 A worked example

Before any abstract countermodels, let us see a concrete case, one that will double as the template for everything that follows. It describes a perfectly ordinary situation: two women, a possible marriage, and a third woman who might have been around but, as things stand, is not.

Take three worlds (see Figure 1), $W = \{w_0, w_1, w_2\}$, with w_0 able to see both w_1 and w_2, so w_0Rw_1 and w_0Rw_2. At w_0 the world contains just Amara and Mei, $D_{w_0} = \{a, m\}$, and they love each other. The world w_1 keeps the same two people, $D_{w_1} = \{a, m\}$ and the love between both. The world w_2 contains a third woman, Soraya, so $D_{w_2} = \{a, m, s\}$, and here it is Mei and Soraya who love each other, while Amara looks on. Writing L for *loves* and using a two-place reading, the interpretation is

$$I(L, w_0) = \{\langle a, m\rangle, \langle m, a\rangle\}, \quad I(L, w_1) = \{\langle a, m\rangle, \langle m, a\rangle\}, \quad I(L, w_2) = \{\langle m, s\rangle, \langle s, m\rangle\}.$$

Assuming all the worlds are reflexive, at w_0 we ask (1) if there are two people who love each other $(\exists x\, \exists y\, (L(x, y) \wedge L(y, x)))$, (2) if there must be $(\Box\, \exists x\, \exists y\, (L(x, y) \wedge L(y, x)))$, and (3) if there is some fixed pair who love each other no matter what $(\exists x\, \exists y \Box\, (L(x, y) \wedge L(y, x)))$.

The first question is settled at w_0 alone. Taking $a, m \in D_{w_0}$ as witnesses, $\langle a, m\rangle \in I(L, w_0)$ and $\langle m, a\rangle \in I(L, w_0)$, so

$$\mathfrak{M}, w_0 \Vdash \exists x\, \exists y\, (L(x, y) \wedge L(y, x)).$$

[1] Each constraint taken singly names a system too, just as in propositional modal logic: reflexivity alone gives **QT**, transitivity alone **QK4**, symmetry alone **QKB**.

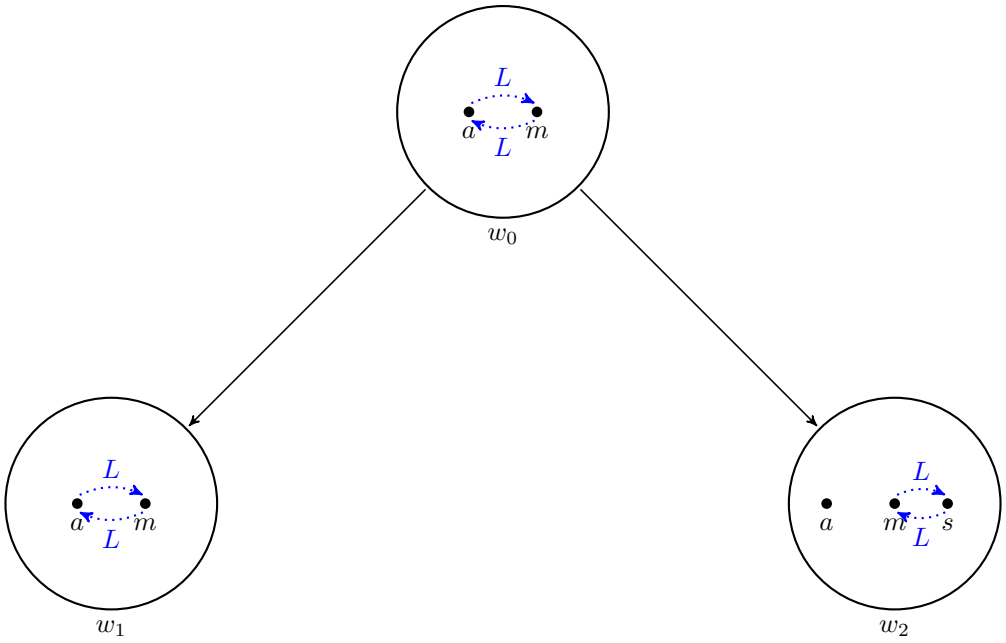


Figure 1: Amara, Mei, and a possible Soraya. Every accessible world contains a loving couple, so the *de dicto* reading holds at w_0. But no couple drawn from $D_{w_0} = \{a, m\}$ loves throughout, because at w_2 the couple is Mei and Soraya, and Soraya does not exist at w_0. The *de re* reading fails.

The second question, the *de dicto* one, asks whether $\Box\,\exists x\,\exists y\,(L(x,y) \wedge L(y,x))$ holds at w_0, that is, whether every accessible world contains some loving pair. At w_1 the pair a, m serves; at w_2 the pair m, s serves. Each accessible world has its own couple, so the box holds:

$$\mathfrak{M}, w_0 \Vdash \Box\,\exists x\,\exists y\,(L(x,y) \wedge L(y,x)).$$

The third question, the *de re* one, is the interesting one. It asks whether $\exists x\,\exists y\,\Box\,(L(x,y) \wedge L(y,x))$ holds at w_0: is there a *single* pair, chosen from D_{w_0}, who love each other at every accessible world? The only candidates are a and m, since they are the only residents of D_{w_0}. They love each other at w_1, but at w_2 we have $\langle a, m\rangle \notin I(L, w_2)$: there, Amara and Mei are not the couple, Mei and Soraya are. So the pair a, m fails the modal tour, and no other pair from D_{w_0} is available to try. Hence

$$\mathfrak{M}, w_0 \nVdash \exists x\,\exists y\,\Box\,(L(x,y) \wedge L(y,x)).$$

The witnessing couple at w_2 was Mei and Soraya, and Soraya simply is not among the individuals at w_0 for the outer quantifier to reach. The *de dicto* reading is satisfied by letting the witness vary from world to world; the de re reading is not, because it must fix the witness in advance, drawing only on what exists here.

2.3.2 The Barcan formulae and their failure

The naïve approach was built on the thought that different worlds may contain different things, and we now reach the first and most important consequence of taking that thought

seriously: two very important principles about the interaction between quantifiers and modal operators fail. They are formulated as follows:

$$\forall x \Box \phi \to \Box \forall x \, \phi \qquad \text{(BF)},$$

$$\Box \forall x \, \phi \to \forall x \Box \phi \qquad \text{(CBF)}.$$

The first, the Barcan formula, states that a universally quantified *de re* statement implies the *de dicto* form of the same universally quantified statement. The second, its converse, states the reverse: that a *de dicto* statement about a universally quantified condition implies the universally quantified *de re* statement. In each the quantifier and the modal operator trade places, in one direction or the other, i.e. $\Box$ and quantifiers *commute*. These formula were first studied by Ruth Barcan Marcus (1946, 1947), who studied syntactic systems of first-order modal logic where both formulas hold.

However, in the current setting both principles fail, and understanding why is key. It is the characteristic signature of varying domains, and seeing exactly why each fails is seeing the approach for what it is. The two fail for opposite reasons, and the reason is exactly the freedom we granted the domains to grow and to shrink.

Take the Barcan formula (BF) first. Its antecedent $\forall x \Box \phi$ quantifies over D_w, the individuals present at the world of evaluation, and says each of them satisfies ϕ at every accessible world. Its consequent $\Box \forall x \, \phi$ moves first to an accessible world v and there quantifies over D_v. If D_v contains a *new* individual, one not among those in D_w, the antecedent has said nothing about it, and there is no reason for it to satisfy ϕ. So the antecedent may hold while the consequent fails. Domain *growth* refutes BF:

Proposition 1. *The Barcan formula $\forall x \Box \phi \to \Box \forall x \, \phi$ is not valid on the class of frames with varying domains.*

Proof. Let $W = \{w, v\}$ with wRv, and let the domain grow: $D_w = \{a\}$, $D_v = \{a, b\}$. Interpret F by $I(F, w) = \{a\}$ and $I(F, v) = \{a\}$, so the newcomer b is not F at v. We check the two sides at w.

For the antecedent $\forall x \Box Fx$: the quantifier ranges over $D_w = \{a\}$, so we need only that a satisfies $\Box Fx$. The sole accessible world is v, and $a \in I(F, v)$, so a is F at v; hence $\mathfrak{M}, w \Vdash \forall x \Box Fx$.

For the consequent $\Box \forall x \, Fx$: we need $\forall x \, Fx$ at v, where the quantifier now ranges over the larger $D_v = \{a, b\}$. But $b \in D_v$ and $b \notin I(F, v)$, so $\mathfrak{M}, v \nVdash \forall x \, Fx$, and therefore $\mathfrak{M}, w \nVdash \Box \forall x \, Fx$.

The antecedent holds and the consequent fails, so (BF) is refuted. The single new individual b, present at v but not at w. See Figure 2. $\Box$

Now the converse formula (CBF). Its antecedent $\Box \forall x \, \phi$ says that at every accessible world v, everything *in* D_v satisfies ϕ. Its consequent $\forall x \Box \phi$ quantifies over D_w and demands that each individual present *here* satisfy ϕ at every accessible world. But an individual $d \in D_w$ may be absent from some accessible D_v. The antecedent constrained only the residents of D_v, and said nothing about our visitor d; under the permissive truth clause d simply fails ϕ there, having no tuple in the extension. So again the antecedent may hold while the consequent fails. Domain *shrinkage* refutes CBF.



Figure 2: Failure of (BF) through domain growth. The antecedent $\forall x \Box Fx$ holds at w because a, the only individual there, is F at the accessible world. But the consequent fails: at v the domain has gained b, which is not F, so $\forall x\, Fx$ is false at v.

Proposition 2. *The converse Barcan formula $\Box \forall x\, \phi \to \forall x \Box \phi$ is not valid on the class of frames with varying domains.*

Proof. Let $W = \{w, v\}$ with wRv, and let the domain shrink: $D_w = \{a, b\}$, $D_v = \{a\}$. Interpret F by $I(F, v) = \{a\}$, and let $I(F, w)$ be anything, say $\{a, b\}$. We check the two sides at w.

For the antecedent $\Box \forall x\, Fx$: we need $\forall x\, Fx$ at v, where the quantifier ranges over $D_v = \{a\}$. Since $a \in I(F, v)$, every resident of D_v is F, so $\mathfrak{M}, v \Vdash \forall x\, Fx$, and hence $\mathfrak{M}, w \Vdash \Box \forall x\, Fx$.

For the consequent $\forall x \Box Fx$: the quantifier ranges over $D_w = \{a, b\}$, so we need each of a and b to satisfy $\Box Fx$, that is, to be F at the accessible world v. The individual b has departed: $b \notin D_v$, and $\langle b \rangle \notin I(F, v)$, so b is not F at v. Thus b fails $\Box Fx$, and $\mathfrak{M}, w \nVdash \forall x \Box Fx$.

The antecedent holds and the consequent fails, so (CBF) is refuted. The departed individual b, present at w but absent at v, is exactly the case the antecedent was silent about. See Figure 3. □



Figure 3: Failure of (CBF) through domain shrinkage. The antecedent $\Box \forall x\, Fx$ holds because at v the only resident a is F. But the consequent fails: b, present at w, is gone at v and so is not F there, so b does not satisfy $\Box Fx$.

So both Barcan formulae fail, BF through worlds that gain individuals, CBF through worlds that lose them. The symmetry is exact, and it is worth dwelling on, because it locates the failure precisely in the freedom we granted the domains.

It is natural to ask whether the failures can be repaired by restricting how the domains may vary, and indeed they can, one at a time. If we forbid the domains to grow, requiring $wRv \Rightarrow D_v \subseteq D_w$, the BF countermodel is blocked and BF becomes valid. If instead we forbid them to shrink, requiring $wRv \Rightarrow D_w \subseteq D_v$, the CBF countermodel is blocked and

CBF becomes valid. Each Barcan direction corresponds to a one-sided inclusion on the domains. But notice what such a repair costs. To restrict the variation in either direction is to give back part of the very freedom that motivated varying domains in the first place; and to recover *both* formulae at once we should have to impose both inclusions together, $D_w \subseteq D_v$ and $D_v \subseteq D_w$, which is to say $D_w = D_v$ across accessible worlds, a single constant domain. At that point the domains no longer vary at all, and the whole naïve picture, with its worlds free to differ in what they contain, has been abandoned. The Barcan formulae can be bought back, then, but only by surrendering the idea that gave the naïve approach its appeal. Whether that surrender is a loss or a liberation is exactly the question on which the two responses of the coming sections divide.

2.3.3 Identity and the necessity of identity

Amid these failures it is worth noticing something that *survives*, for it will matter later and it shows that varying domains do not wreck everything. Add an identity symbol, interpreted at each world as genuine identity on that world's domain:

$$I(=, w) = \{\langle d, d\rangle \mid d \in D_w\}.$$

Because constants are rigid, denoting one fixed individual at every world, the necessity of identity goes through much as in the constant-domain case, provided the individuals in question exist where we evaluate.

Proposition 3. *For constants b, c, the formula $b = c \to \Box(b = c)$ is valid on every frame with varying domains, at any world from which b and c continue to denote existing individuals throughout the accessible worlds.*

Proof. Suppose $\mathfrak{M}, w \Vdash b = c$. Then $\langle I(b), I(c)\rangle \in I(=, w)$, so $I(b) = I(c)$; call this individual d. By rigidity, $I(b)$ and $I(c)$ are this same d at every world. Let v be accessible from w. Wherever $d \in D_v$, we have $\langle I(b), I(c)\rangle = \langle d, d\rangle \in I(=, v)$, so $\mathfrak{M}, v \Vdash b = c$. As v was arbitrary among the accessible worlds in which d exists, $\mathfrak{M}, w \Vdash \Box(b = c)$. □

This is again the formal counterpart of Kripke's thesis from Section 1.3: co-designating names co-designate of necessity, and the contingency of how we *discover* an identity leaves no mark on the formula. *Hesperus is Phosphorus*, once true, stays true at every accessible world in which the planet exists, and the astronomical labour of its discovery is invisible to the logic.

The same pressure bears on the necessity of *distinctness*, $b \neq c \to \Box(b \neq c)$. Where both individuals exist it holds for the same reasons; but the moment one of them goes missing, the permissive clause makes the matter turn on a technicality about absent objects rather than on the standing fact of their being two. We let these delicacies stand for now. They are not defects of the naïve approach so much as symptoms of the one decision in its semantics that was made for convenience rather than principle, the decision to count predications about absent objects false. Taking that decision seriously, and asking what a more principled treatment of non-existence would look like, is precisely the project of the section on contingentism, where free logic is built to handle exactly these cases.

2.4 The Failure of Completeness, Intuitively †

We have seen what the naïve system does well, and we have seen the metaphysical discomfort of the Barcan failures. There is a further trouble, of a quite different kind. It is not a matter of which formulae come out valid, but of whether the logic can be matched, as a calculus, to its own semantics. For the systems we care about most, it cannot, at least not by the natural means, and the reason is worth seeing because it traces back to the very freedom that defines the approach.

A logic has two sides. On one side is its semantics: the models, and the formulas true in all of them, which we call *valid*. On the other is its syntax/proof system: some axioms and rules, and the formulas derivable from them, which we call *theorems*. These answer to two different ambitions. The valid formulas are those genuinely true under the intended reading of the operators; the theorems are those we can actually establish by reasoning. A logic is *sound* when every theorem is valid, so the proof system never leads us astray, and *complete* when, conversely, every valid formula is a theorem, so the proof system is strong enough to reach every genuine validity.

For propositional modal logic, the two faces match across the board. The system **K** is complete for the class of all frames, **S5** for the frames whose accessibility relation is an equivalence relation, and likewise for the systems between. The valid and the derivable coincide. One naturally expects this to survive the passage to the quantified setting, since we built the quantified systems by adding the predicate-logical apparatus, itself complete, on top of these complete propositional bases. Surely the combination inherits the completeness of its parts. For many of systems it does, but not for all of them.

Before we get our hands dirty with the machinery of canonical models, let us build some intuition for what goes wrong. If you have never seen a completeness proof, do not worry: the moral of this subsection can be grasped without one, and it is the moral, not the machinery, that matters for the rest of the chapter.

Soundness and completeness are two halves of a promise. Soundness promises that the calculus never lies: everything you can *prove* is genuinely *true* in every model. Completeness promises the converse: every formula *true* in every model can actually be *proved*. A logic for which both hold is one where "valid" and "provable" pick out exactly the same formulae, so we may reason with whichever notion is convenient without ever being led astray.

The standard way to *deliver* on the completeness half is disarmingly direct. Suppose someone hands us a formula ϕ we *cannot* prove. To show the calculus is complete we must exhibit a single model in which ϕ actually fails. If we can always produce such a counter-model on demand, nothing slips through: every unprovable formula is refuted somewhere, so every formula true *everywhere* is provable. Completeness, then, is a promise that *we can always build the model we need*.

Where do we find such a model? The ingenious idea, due to Henkin, is to build it out of the syntax itself. Rather than search the space of models hoping to stumble on a suitable one, we assemble a bespoke model whose *worlds are maximal consistent sets of sentences*: bundles of formulae that are internally coherent and that have settled every sentence of the language one way or the other. A world, on this construction, is nothing more than a complete and consistent description of how things stand. One world is deemed to *access* another exactly when everything the first holds to be necessary is held outright by the second; and each world carries a domain assembled from the terms that world's bundle

certifies as denoting. This is the **canonical model**, and its whole purpose is captured by the **Truth Lemma**: at each world a sentence is *true* precisely when it is a *member* of that world's bundle. Truth and membership march in lockstep.

If the Truth Lemma holds, completeness falls out at once. An unprovable ϕ leaves $\neg\phi$ consistent; extend $\neg\phi$ to one of our bundles; that bundle is a world of the canonical model where, by the Lemma, $\neg\phi$ is true and hence ϕ is false. The countermodel we were promised has built itself out of the syntax.

But there is a subtlety we have so far suppressed. The canonical model is one *fixed* object: its worlds, its domains, and above all its accessibility relation are all read off the syntax, with no further choices left to us. For the countermodel to count, this canonical accessibility relation must have whatever shape the logic requires. If our logic is **QS5**, so that $\Box$ is to mean truth in *every* possibility without restriction, the relation must come out an equivalence relation, and in particular *symmetric*: whenever a world w can see a world v, the world v must be able to see w in return. Completeness for **QS5** is thus a promise about the canonical model specifically, that the relation the construction hands us is symmetric. The failure of the naïve approach is that, for varying domains, we cannot keep this promise.

To see why, we watch the construction try to honour symmetry and jam. The mechanism is exactly the one behind the Barcan failures: a quantifier inside a box and a quantifier outside it answer to *different* domains, and the syntax gives us no way to make the two agree across an accessible pair.

Fix a world Γ of the canonical model, and suppose Γ contains $\exists x\, Px$, "something is P". A maximal consistent bundle does not merely record *that* something is P; being complete, it must be saturated (i.e.*names* a witness), committing to some particular term c with $Pc \in \Gamma$. This is the analogue, at the level of syntax, of choosing a resident of D_Γ to witness the existential: the witness is fixed *here*, at Γ, drawn from what Γ acknowledges to exist.

Now let Γ access a world Δ. Because the domains vary, Δ may certify individuals of its own that Γ never named, and may decline to certify Γ's witness c. Suppose Δ commits to $\Box Pd$ for some *new* term d, one that denotes at Δ but not at Γ. If the relation is to be symmetric, Δ must access Γ in return; and then the accessibility clause demands that everything Δ holds necessary holds also at Γ, so from $\Box Pd \in \Delta$ we should need $Pd \in \Gamma$. But Γ never certified d at all: d is one of Δ's own residents, a merely possible individual from Γ's point of view, and Γ committed only to its *own* witness c. The witness Γ fixed downstream and the witness Δ demands upstream need not be the same individual, and with varying domains there is nothing to make them coincide. We have represented the situation just described in Figure 4.

So the backward link required by symmetry cannot be forged. The construction delivers a canonical accessibility relation, but not a symmetric one: the witnesses each world chooses, drawn from its own domain, are chosen independently, and the domains do not line up across the pair. The Truth Lemma still holds world by world, and yet the model it certifies is not an **S5** model, because its relation is not symmetric. Our one recipe for manufacturing countermodels produces, for **QS5**, a model of the wrong shape.

It is worth being precise about what has and has not gone wrong, since the situation is more delicate than outright unsoundness or a broken axiom. Nothing here shows that **QS5** has no complete calculus; with more sophisticated constructions it can be shown complete after all. What fails is the *transfer* of completeness from the propositional base

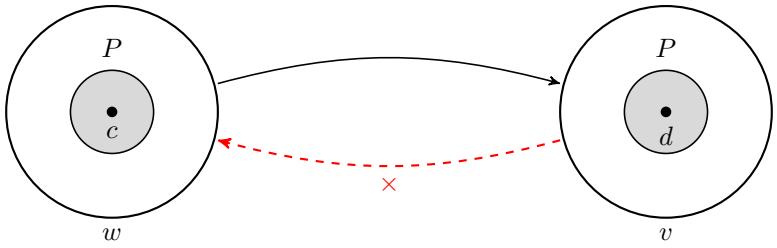


Figure 4: The breakdown. In the canonical model for a symmetric system, the world w (playing Γ) chooses its own witness c for $\exists x\, Px$, drawn from what it acknowledges to exist. The accessed world v (playing Δ) settles its own affairs and declares $\Box Pd$ for an individual d of its own, one that does not exist at w. Symmetry would demand a return arrow $v \to w$ carrying the necessity $\Box Pd$ back, so that Pd must hold at w; but $d \notin D_w$, and the witnesses fixed on either side never coincide. The return arrow cannot be honoured, and the canonical relation fails to be symmetric. This is the clash the constant-domain repair of Figure 5 dissolves.

to its quantified counterpart by the *canonical* route. In the propositional case the canonical relation for **S5** comes out symmetric for free, and completeness is immediate. Adding varying domains breaks the guarantee: the very freedom that let worlds gain and lose individuals, the freedom that refuted the Barcan formulae, now lets a world's chosen witnesses drift out of an accessing world's reach, and the canonical relation is no longer forced into the shape the logic needs.

The pattern is entirely general, and it sorts the systems cleanly. The construction still works for logics whose frame conditions look only *forward* along the accessibility relation: reflexivity, transitivity, and the like are inherited by the canonical relation automatically, so **QK**, **QT**, **QK4** and **QS4** are complete by exactly this route. It breaks for logics whose conditions look *backward*, demanding that a link be answered by a link in the reverse direction: symmetry, and with it the euclidean condition, are not secured by the syntax, and **QKB** and **QS5** slip through. The dividing line is precisely forward-looking versus backward-looking axioms, and it is drawn by the same varying domains that drive everything else in this chapter.

Two ways out of incompleteness present themselves, and they divide the road ahead. One may refuse to let the domains vary at all, returning to a single constant domain shared by every world, restoring the Barcan formulae and the canonical construction together, and paying the metaphysical price of holding that whatever exists exists necessarily. This is the road of *necessitism*, and we take it up with an identity predicate in the next chapter. Or one may keep the varying domains, embracing the contingency of existence they express, and instead rebuild the logic from below, replacing its classical treatment of the terms that fail to denote with a free logic honest about non-existence. This is the road of *contingentism*. The completeness failure is the fork itself: the naïve approach cannot stand, and the two chapters that follow are the two ways of leaving it.

Exercises

Exercise 5

Let $\mathfrak{M}$ have $W = \{w_1, w_2\}$ with $w_1 R w_2$ and $w_2 R w_2$ as the only arrows, domains $D_{w_1} =$

$\{a\}$ and $D_{w_2} = \{a, b\}$, and $I(F, w_1) = \{a\}$, $I(F, w_2) = \{b\}$. Compute the truth value at w_1 of each formula, showing your working in the manner of the text.

(a) $\Box \forall x\, Fx$

(b) $\forall x \Box Fx$

(c) $\Box \exists x\, Fx$

(d) $\exists x \Box Fx$

(e) $\Diamond \exists x \neg Fx$

Exercise 6

Using the permissive truth clause for atomic formulas, evaluate the following at the world w of the model $W = \{w, v\}$, wRv, with $D_w = \{a, b\}$, $D_v = \{a\}$, and $I(P, w) = \{a\}$, $I(P, v) = \{a\}$. In each case say which clause forces the value, and pay particular attention to predications about b at v, where b does not exist.

(a) $\Box P(b)$

(b) $\Diamond P(b)$

(c) $\forall x \Box P(x)$

(d) $\Box \forall x\, P(x)$

Exercise 7

Build a countermodel, drawn as a labelled diagram, for each of the following. State the frame, the domains, and the interpretation explicitly, and identify the world at which the formula fails.

(a) $\forall x \Box Fx \to \Box \forall x\, Fx$ (BF)

(b) $\Box \forall x\, Fx \to \forall x \Box Fx$ (CBF)

(c) $\Diamond \exists x\, Fx \to \exists x \Diamond Fx$

For each, say in one sentence whether the countermodel turns on a domain that *grows* or one that *shrinks* along the accessibility relation.

Exercise 8

Section 2.3 observed that BF can be reinstated by forbidding the domains to grow, and CBF by forbidding them to shrink. Confirm both halves semantically.

(a) Prove that if a model satisfies $wRv \Rightarrow D_v \subseteq D_w$ at every pair of worlds, then BF is valid in it. Where does your argument use the inclusion?

(b) Prove that if a model satisfies $wRv \Rightarrow D_w \subseteq D_v$ at every pair, then CBF is valid in it.

(c) Conclude that a model validating both BF and CBF on every accessible pair has $D_w = D_v$ whenever wRv.

Exercise 9*

(a) The schema $\Box\phi \to \phi$ is the reflexivity axiom. Build a varying-domain countermodel to its instance $\Box\forall x\, Fx \to \forall x\, Fx$ by choosing a frame in which some world has no reflexive arrow. Conclude that this instance is not valid in **QK**.

(b) The schema $\Box\phi \to \Box\Box\phi$ is the transitivity axiom. Build a three-world countermodel to an instance of it by choosing R with w_1Rw_2 and w_2Rw_3 but *not* w_1Rw_3. (You may use a purely propositional ϕ, or a quantified one; say which features of the domains, if any, your countermodel uses.)

(c) In each of (a) and (b), the formula *becomes* valid once the appropriate condition is imposed on R. Explain, in a sentence each, why imposing the condition blocks your countermodel. This is the sense in which the validities depend on the choice of frame class, not on the truth definition alone.

Exercise 10

Inspector Mehta is interviewing three suspects, and the case admits several possibilities, modelled as worlds accessible from the actual world w_0. Let G be the predicate *is guilty*. Translate each statement into the language of Section 2, then say whether, on the intended reading, it is a *de dicto* or a *de re* claim, and construct a small model on which the two come apart.

(a) "It is possible that everyone present is innocent."

(b) "Everyone present might be the guilty one."

(c) "Necessarily, someone present is guilty."

(d) "There is someone present who is necessarily guilty."

Exercise 11

Return to the model of Amara, Mei, and a possible Soraya (Figure 1). Without changing the frame or the domains, alter only the interpretation of L so that, at w_0:

(a) the de re formula $\exists x\,\exists y\,\Box(L(x,y) \wedge L(y,x))$ becomes *true*;

(b) both the de dicto and the de re readings become false;

(c) the de re reading holds but the de dicto reading fails. (Is this possible? If not, prove that it is not.)

Exercise 12*

We add an identity symbol, interpreted as in Section 2.3, $I(=, w) = \{\langle d, d\rangle : d \in D_w\}$.

(a) Construct a two-world model and constants b, c for which $b = c \to \Box(b = c)$ *fails*, by letting the individual they denote drop out of the domain of an accessible world. Identify precisely which clause makes the identity false at that world.

(b) Show that if the denoted individual exists at every accessible world, the same formula holds. Where does rigidity of the constants enter?

(c) Discuss, in two or three sentences, whether the failure in (a) is a genuine failure of the necessity of identity or an artefact of the permissive truth clause.

Exercise 13**

The failure subsection argued that the natural model-building construction, run on a symmetric system, produces a relation that is not symmetric, because domains grow forward and cannot be pushed back. This exercise asks you to localise that argument in a single small picture.

Consider building a model in which a world w realises a possibility at an accessible world v, and v in turn names a fresh individual $d \notin D_w$ and declares $\Box P(d)$.

(a) Write down what symmetry would require to hold back at w, and explain why the construction cannot supply it.

(b) Suppose, contrary to the construction, that we *stipulate* $D_w = D_v$ from the outset. Show that the obstruction in (a) disappears.

(c) Relate your answer to (b) to Exercise 4(c): explain in what sense "hold the domains constant" is simultaneously the repair for the Barcan failures and the repair for the broken construction. (This is the convergence the next section builds on.)

3 The First Response: Necessitism

The previous section ended in a genuine difficulty. The naïve approach, with its freely varying domains, refutes both Barcan formulae and, worse, resists the one construction that would tie it to a proof system: run on a symmetric logic, the natural method builds a model whose accessibility relation is not symmetric, and so fails to model the very logic it was meant to capture. The route from the proof system of **QS5** to a matching class of models, the route that ran so smoothly for the forward-looking systems, was blocked. Something must give.

There are, broadly, two ways to respond, and they part company over which feature of the naïve picture to blame. One can let the varying domains stand, accept that the Barcan formulae fail, and look for a logic that lives comfortably with worlds of differing population; this is the road to contingentism, and we take it in Section 4. Or one can locate the trouble in the variation itself, and remove it: hold the domain fixed across all worlds, so that nothing is ever gained or lost, and watch both the Barcan formulae and the broken construction come right at once. This second road is the one we follow here. It is the more conservative of the two, in that it restores the clean classical behaviour the naïve approach disturbed, which serves to ground a philosophical position. Namely *necessitism*,

the thesis that necessarily everything necessarily exists –that the inventory of the world could not have been otherwise. The position has a forceful contemporary defender in Timothy Williamson (2013), who argues that the constant domain is not a convenient idealisation to be explained away but the correct account of how being and necessity interlock. We will impose the constant domain, see how it repairs both ailments, sharpen the repair with a small and elegant piece of structure, and then face squarely the commitment it carries.

3.1 The Fix: Hold the Domain Fixed

The diagnosis of the previous section pointed, twice over, in the same direction. The Barcan formulae failed because domains were free to grow and to shrink; we noted there that forbidding growth restores one formula, forbidding shrinkage the other, and forbidding both, that is, requiring $D_w = D_v$ whenever wRv, restores them together. A domain that neither grows nor shrinks is exactly what would let a forward arrow be matched by a backward one. So a single stipulation answers both complaints.

Definition 7 (Constant Domain)**.** *A frame $\langle W, R, D\rangle$ has* constant domains *if $D_w = D_v$ for all $w, v \in W$. Equivalently, there is a single domain D, shared by every world, over which every quantifier ranges wherever it is evaluated.*

With the domain held fixed, the semantics of Section 2.2 simplifies in one stroke. The awkward case that forced the permissive truth clause, a variable denoting an object absent from the world of evaluation, can no longer arise, since no object is ever absent: every individual exists at every world. The quantifier $\forall x$ ranges over the same D at w and at every accessible v, so the disagreement between box and quantifier that drove the whole of the previous section simply vanishes. We record the immediate payoff.

Proposition 4. *On every constant-domain frame, both Barcan formulae are valid:*

$$\forall x\,\Box\phi \to \Box\,\forall x\,\phi \quad \text{(BF)}, \qquad \Box\,\forall x\,\phi \to \forall x\,\Box\phi \quad \text{(CBF)}.$$

Proof. For (BF), suppose $\mathfrak{M}, w \Vdash_g \forall x\,\Box\phi$, i.e. for every $d \in D$ we have $\mathfrak{M}, v \Vdash_{g[x\mapsto d]} \phi$ at every v with wRv. Fix such a v. To show $\mathfrak{M}, v \Vdash_g \forall x\,\phi$ we must check $\mathfrak{M}, v \Vdash_{g[x\mapsto d]} \phi$ for every $d \in D_v$. But $D_v = D$, the very set the antecedent ranged over, so this is exactly what we have. Hence $\Box\,\forall x\,\phi$ holds at w.

For (CBF), suppose $\mathfrak{M}, w \Vdash_g \Box\,\forall x\,\phi$, so that $\mathfrak{M}, v \Vdash_g \forall x\,\phi$ at every accessible v. Take any $d \in D$ and any accessible v. Since $D_v = D$, $d \in D_v$, and $\forall x\,\phi$ holds at v, we get $\mathfrak{M}, v \Vdash_{g[x\mapsto d]} \phi$. As v was arbitrary, $\mathfrak{M}, w \Vdash_{g[x\mapsto d]} \Box\phi$; and as d was arbitrary, $\forall x\,\Box\phi$ holds at w. □

Each direction uses the fact that an individual quantified over at one world is automatically available to be quantified over at any other. This is the move the varying-domain semantics withheld and the constant domain now grants. The two formulae that the naïve approach refuted are, under constancy, theorems; box and quantifier commute without remainder.

That settles the Barcan formulae. We will now turn to a further refinement that makes the metaphysical picture even more clear.

3.2 A Refinement: Kripke Models with Equality

The structure we are after is a relation of *equality* on the domain, world by world, and although a determined logician could do without it, we introduce it because it shows, in a single picture, exactly how constancy mends the broken construction, and because it gives the necessity of identity a home in the frame itself.

Recall precisely how the construction broke. A world w chose a witness c for every existential formula; an accessible world v delivered a necessity $\Box Pd$ about some *other* individual d; symmetry demanded that Pd hold back at w; but w knew only its own witness c, and c and d were simply two different names with no bridge between them. What the construction lacked was any way for w, looking back along the return arrow, to *recognise* the required individual d as one it already possessed. A constant domain guarantees that d *is* present at w, since every object exists at every world; but presence under a possibly different name is not yet recognition. The equality relation supplies the recognition.

Definition 8 (Kripke Model with Equality)**.** *A* Kripke model with equality *is a tuple* $\mathfrak{M} = \langle W, R, D, \asymp, I\rangle$ *where* $\langle W, R, D\rangle$ *is a constant-domain frame,* I *interprets each* n*-ary predicate* P *at each world as an extension* $I(P, w) \subseteq D^n$*, and* $\asymp = \{\asymp_w\}_{w\in W}$ *assigns to each world* w *an equivalence relation* $\asymp_w$ *on* D*, the* equality relation *at* w*, subject to:*

- ***Heredity:*** *if* wRv *and* $a \asymp_w b$*, then* $a \asymp_v b$*. Equality once recognised is never undone as we move to accessible worlds.*
- ***Congruence:*** *if* $a \asymp_w b$*, then* a *and* b *are interchangeable in* $I(P, w)$ *for every* P*; equal objects satisfy the same predicates at the world where they are equal.*

The truth clause for identity is at follows

$$\mathfrak{M}, w \Vdash_g x = y \quad \Longleftrightarrow \quad g(x) \asymp_w g(y),$$

and all other clauses are exactly those of the constant-domain semantics.[2]

Formally, $a \asymp_w b$ says that the logic is entitled to treat a and b as one and the same individual at w, even should they be distinct points of the underlying set D. This is the slack the naïve construction lacked: a world may now carry an object that is, to the logic at that world, indistinguishable from another, so that a witness chosen under one name can answer to a demand framed under another. Read metaphysically, in the necessitist key, $a \asymp_w b$ says that a and b *are* the same individual at w, the domain having been stocked once and for all with objects that the equality relation then organises into individuals.

It is worth checking that nothing we cared about is disturbed. The quantifier still ranges over the one fixed D and the box still carries the assignment unchanged across accessible worlds, so Proposition 4 stands: both Barcan formulae remain valid. What the equality relation adds is read entirely in the identity statements, and there it adds exactly the structure the repaired construction needs.

3.3 Reclaiming Completeness†

We can now say, without grinding through the construction in full, why the addition repairs what the naïve approach could not. The completeness proof is once more the canonical

[2] For more on Kripke structures with equality, see (Gabbay et al., 2009, §3.5).

model, the model read off the logic's own consistent theories; but the objects of the model are no longer names. Two names are taken to denote the same individual whenever the logic proves them equal, and the equality relation of the canonical model is precisely this: $a \asymp_w b$ holds when the theory at w contains the equation between them. The identity axioms make this an equivalence relation and a congruence, so the canonical model is a genuine Kripke model with equality in the sense of Definition 8.

The decisive step that broke before is now mended. When symmetry forces a return arrow, and some claim about d must hold at world, the world need no longer contain the very name d. It need only contain an object *equal* to d, and the equality relation gives it the means to look back and recognise that it does. The witness chosen going forward and the object demanded coming back are reconciled, not by being the same name, but by being equal. Heredity ensures that this recognition, once available, is never lost as we move about the frame, so the construction is stable; and because the domain is constant, the object the return arrow demands is guaranteed to be present to be recognised in the first place. The mismatch that had no resolution under varying domains now has one, supplied jointly by the constant domain (which guarantees presence) and the equality relation (which supplies recognition).

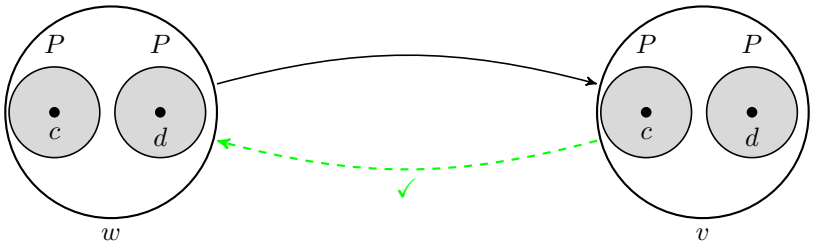


Figure 5: The repair. With a constant domain the individual d is present at w to begin with, and the equality relation lets w recognise its own witness c as equal to the returning d. The demand Pd that broke the naïve construction (cf. Figure 4) is now met, and the return arrow can be honoured.

Notice, that constant domain alone already validates the Barcan formulae, and completeness for **QS5** can be obtained, by a more laborious route, without ever naming an equality relation as part of the model. We have introduced it because it earns its keep twice over. It makes the mechanism of the repair visible, turning a delicate argument about accumulating witnesses into a single picture of a world recognising one of its own; and, as we now see, it hands the necessitist a clean account of identity as a bonus.

3.4 The Commitment: Necessitism

It would be comfortable to present the foregoing as a purely technical repair with no metaphysical residue. It is not. We imposed a constant domain, we thereby validated both Barcan formulae, and in reading the equality relation as we did we took a substantive stand on what there is.

Consider what the two Barcan formulae jointly assert once the quantifier is read as ranging over what exists as a single domain. The Barcan formula, in the form $\Diamond \exists x\, \phi \rightarrow \exists x \Diamond \phi$, reads that if there could have been a thing of some kind, then there already is a thing that could have been of that kind: no merely possible individual exist beyond the

actual ones. The Converse, in the form $\exists x \Diamond \phi \to \Diamond \exists x \phi$, states that if there is an object that is possibly of some kind, then it is possible that there is an object of this kind: no actual individual could have failed to exist. Together they make the inventory of the world modally static. Necessarily, everything necessarily exists. This is necessitism, and our semantics, with its single fixed D and its world-invariant equality, is its exact formal image.

Stated baldly, the thesis sounds extravagant, and the objections come at once. Surely this page might never have existed; surely there might have been people other than those there are. The necessitist does not deny the phenomena behind these complaints. What is denied is the inference from them to a *varying domain*. The move, which is Williamson's, is to distinguish *being* from *being concrete*. Everything there is, is of necessity; what is contingent is not whether a thing *is* but whether it is *concrete*, located in space and time and woven into the causal order. On this reading this page does exist at every world, as an object in the one domain D; what varies from world to world is only whether it is there a concrete printed thing or a merely possible, non-concrete one. Coming into being and passing away are then not changes in what there is, but changes in how *what there is* is qualified, changes recorded by the world-dependent predicates while the domain stands fixed. The semantics expresses this with no strain at all: the predicates vary, the domain does not, and the equality relation tells us, stably across all worlds, which objects are one and the same.

Read so, the validity of the Barcan formulae stops looking like an embarrassment and starts looking like a structural truth the logic was right to enforce, and the repaired construction becomes a sign that the logic has correctly captured its subject matter rather than a mere technical convenience. The equality relation wears its two faces at once. As a formal device it is what lets a world look back and recognise a returning object; as metaphysics it is the relation of genuine sameness among the necessarily existing things, modally stable by heredity, which is to say that identity and distinctness hold of necessity. The device and its necessitist reading are, in the end, the same structure seen from two sides.

Exercises

Exercise 14

Work in a constant-domain model $\mathfrak{M} = \langle W, R, D, I \rangle$ (no equality relation needed). Verify the following by direct semantic argument from the truth clauses.

(a) $\Diamond \exists x\, Fx \to \exists x \Diamond Fx$ is valid.

(b) Conclude that, on constant domains, $\Diamond \exists x\, Fx$ and $\exists x \Diamond Fx$ are equivalent, as are $\forall x \Box Fx$ and $\Box \forall x\, Fx$.

Exercise 15

(a) Prove that the converse Barcan formula $\Box \forall x\, \phi \to \forall x \Box \phi$ is valid on a frame whenever the domains never shrink along accessibility, i.e. $wRv \Rightarrow D_w \subseteq D_v$.

(b) Prove that the Barcan formula $\forall x \Box \phi \to \Box \forall x\, \phi$ is valid whenever the domains never grow, i.e. $wRv \Rightarrow D_v \subseteq D_w$.

(c) Conclude that both are valid exactly when the domains are locally constant ($wRv \Rightarrow D_w = D_v$), and explain in a sentence why, on a frame whose R connects every world to every other, local constancy collapses to a single global domain.

Exercise 16

Show *semantically* that the two Barcan formulae imply each other on symmetric frames. That is, let R be symmetric and suppose every instance of (BF) holds on the frame; show every instance of (CBF) holds, and conversely.

(a) For one direction, suppose (CBF) could fail at some w: there is a $d \in D_w$ with $d \notin D_v$ for some accessible v witnessing the failure. Use symmetry to view the situation from v, where w is now accessible, and derive a corresponding failure of (BF) there.

(b) State, in a sentence, what this mutual implication says about the domain inclusions of Exercise 2: why can a symmetric frame not have domains that grow without also having them shrink?

Exercise 17

The necessitist thesis can be written

$$\Box \forall x \, \Box \exists y \, (x = y), \tag{NT}$$

"necessarily, everything is necessarily identical to something". Work in a constant-domain Kripke model with equality.

(a) Show that (NT) is valid on every such model. (Hint: the inner $\exists y \, (x = y)$ asks only that the object assigned to x be $\asymp_w$-equal to *some* resident of D_w; reflexivity of $\asymp_w$ and the constancy of D settle it.)

(b) Explain why the argument breaks down under varying domains: exhibit a two-world model in which (NT) fails, and say which conjunct of the definition of a constant-domain frame your countermodel violates.

(c) In one or two sentences, connect (a) and (b) to the metaphysical reading of Section 3.4: in what sense is the validity of (NT) the formal content of "nothing could have failed to be"?

Exercise 18*

Show that the necessity of identity,

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

follows from the necessitist thesis together with the structure of a Kripke model with equality. You may argue either semantically or syntactically.

(a) *Semantically:* suppose $\mathfrak{M}, w \Vdash_g a = b$, so $g(a) \asymp_w g(b)$. Use the heredity condition on $\asymp$ to push the equality forward to every accessible world, and conclude $\mathfrak{M}, w \Vdash_g \Box(a = b)$. Identify the exact role heredity plays.

(b) *Syntactically:* take (NT) and the converse Barcan formula as available, and derive the necessity of identity. (You may find it cleanest to reason about $\exists y\,(x = y)$ and substitute.)

(c) Which of the two arguments makes plainer *why* the result holds, the semantic one resting on heredity or the syntactic one resting on (CBF)? Defend your answer in a sentence.

Exercise 19*

The equality relation, we said, is a refinement that is not strictly indispensable. This exercise probes what it adds.

(a) Consider the *congruence* condition (equal objects satisfy the same predicates at the world where they are equal). Construct a tuple $\langle W, R, D, \asymp, I\rangle$ satisfying every clause of Definition 8 *except* congruence, and exhibit a formula that comes out true under one name of an object and false under an $\asymp_w$-equal name. Conclude that congruence is doing genuine work.

(b) Consider *heredity*. Construct a tuple satisfying every clause except heredity, with $a \asymp_w b$ but $a \not\asymp_v b$ for some accessible v, and show that the necessity of identity $a = b \to \Box(a = b)$ then fails. Relate this to the metaphysical claim that, for the necessitist, sameness is modally stable.

Exercise 20**

This exercise revisits the repair of the broken construction (Section 3.3) from the model-building side, in a single small picture.

Build two worlds w, v with wRv, intended as fragments of a canonical model. At w, an existential is witnessed by a name c. At v, a necessity $\Box P(d)$ is recorded for a name d, and symmetry requires vRw as well.

(a) Under *varying* domains with no equality relation, explain why the return arrow cannot be honoured, recovering the breakdown of Figure 4.

(b) Now impose a constant domain and an equality relation with $c \asymp_w d$. Show that $P(d)$ now holds at w, so the return arrow *can* be honoured. Which of the two additions, the constant domain or the equality relation, guarantees that d is *present* at w, and which guarantees that w *recognises* it?

(c) Suppose we kept the constant domain but dropped the equality relation, reading $=$ as bare identity of names. Does the return arrow close? Argue that it does, and hence that the equality relation, however illuminating, is not what makes completeness possible, only what makes the recognition explicit. (This is the honesty of Section 3.3 turned into an exercise.)

Exercise 21

The necessitist reads the validity of (NT) as the discovery that "nothing could have failed to be," explaining away apparent cases of coming-to-be as changes in concreteness rather than in being (Section 3.4).

(a) State, in your own words, what a defender of varying domains would say is *wrong* with treating the validity of (NT) as a metaphysical discovery rather than an artefact of the constant-domain stipulation.

(b) The necessitist replies that the varying-domain semantics is itself a stipulation, with no better claim to neutrality. Assess, briefly, whether the dispute can be settled by the formal semantics alone, or whether it turns on something the semantics cannot supply.

4 The Second Response: Contingentism

The necessitist met the trouble of the naïve approach by removing what caused it. Domains varied, so the necessitist held them fixed; and with the domains fixed, both Barcan formulae returned as theorems and the broken construction closed. The price was a thesis about being: necessarily, everything necessarily exists. The contingentist takes the opposite view of the very same trouble. Where the necessitist saw the varying domains as the disease, the contingentist sees them as the diagnosis. That the Barcan formulae fail is not a defect to be corrected since it really is contingent what there is. Had your parents never met, you would not have existed; there might have been people other than those there are, and there might have been fewer. A logic that registers this, by letting the domains vary and letting the Barcan formulae lapse, is a logic that has got something right, not something wrong.

So the contingentist keeps the varying domains of Section 2 and looks elsewhere for the repair. And there is a repair to be made, for we left the naïve approach with two debts outstanding. The first was a frankly expedient choice in the truth definition, that a predication about an absent object simply comes out false. The second was the awkwardness of fitting the logic to a proof system at all. The contingentist pays the first debt by reforming the logic of the quantifiers, moving to a *free* logic in which terms may fail to denote and the quantifiers carry existential import only over what actually exists. This is the move that earns contingentism its name, and it carries, like necessitism's move, a cost of its own, here a cost paid in the classical inference rules rather than in metaphysics. The position has a forceful recent defender in Greg Restall (2019, 2026), who develops it not from a model theory at all but from the proof theory of inference, and who turns the usual order of explanation on its head. We begin with his reason for doing so, because it frames everything that follows.

4.1 The Semantics: Varying Domains and Free Logic

The frames are exactly those of the naïve approach: a set of worlds, an accessibility relation, and a domain D_w for each world, with no constraint relating one world's domain to that of another. What changes is the treatment of terms and quantifiers, and the change has a single source: the contingentist takes seriously that a term may denote an object absent from the world of evaluation, and reforms the logic so that absence is handled better.

Two ideas do the work. The first is a single *global* stock of individuals, $D = \bigcup_w D_w$, the objects that exist at some world or other, from which terms take their values; an individual constant denotes one such object rigidly, the same at every world, but that object need not

belong to the local domain D_w of the world where we evaluate. The second is an *existence predicate*, written $E!$, true at w of exactly the residents of D_w:

$$\mathfrak{M}, w \Vdash_g E!t \quad \iff \quad [\![t]\!]^{\mathfrak{M},g} \in D_w.$$

The existence predicate is the contingentist's central instrument. It lets the object language say, of a term, that it denotes something *here*, and it is what the quantifier rules will consult.

Definition 9 (Contingentist Model)**.** *A* contingentist model *is a tuple* $\mathfrak{M} = \langle W, R, D, I\rangle$ *where* $\langle W, R, D\rangle$ *is a frame with varying domains,* $D = \bigcup_{w\in W} D_w$ *is the global domain, and* I *assigns to each constant a fixed individual* $I(c) \in D$ *and to each* n*-ary predicate* P *and world* w *an extension* $I(P, w) \subseteq D^n$ *subject to the* residence condition*:*

$$\langle a_1, \ldots, a_n\rangle \in I(P, w) \quad \Longrightarrow \quad a_1, \ldots, a_n \in D_w.$$

Assignments map variables into the global domain D*.*

Notice the one quiet but decisive constraint in this definition. Although $I(P, w)$ draws its tuples from the *global* domain D, the residence condition forbids any tuple with a non-resident component. This is the choice of a *negative* free logic: a predicate cannot hold at w of an object that fails to exist at w. It is what makes, for example, *Pegasus is a winged horse* come out false at our world.[3]

The truth definition now splits the quantifier from the predicate in the way free logic requires. Predication reads off the extension directly, and the residence condition guarantees that an atomic formula about an absent object simply fails; the quantifier, by contrast, ranges only over the *local* domain.

Definition 10 (Truth, contingentist semantics)**.** *For a contingentist model* $\mathfrak{M}$*, world* w*, and assignment* g*:*

- $\mathfrak{M}, w \Vdash_g P(t_1, \ldots, t_n)$ *iff* $\langle [\![t_1]\!]^{\mathfrak{M},g}, \ldots, [\![t_n]\!]^{\mathfrak{M},g}\rangle \in I(P, w)$*; by the residence condition this requires each* $[\![t_i]\!]^{\mathfrak{M},g} \in D_w$*;*
- $\mathfrak{M}, w \Vdash_g E!t$ *iff* $[\![t]\!]^{\mathfrak{M},g} \in D_w$*;*
- *the clauses for* $\neg$*,* $\to$*, and* $\Box$ *are as before;*
- $\mathfrak{M}, w \Vdash_g \forall x\, \phi$ *iff for every* $d \in D_w$*,* $\mathfrak{M}, w \Vdash_{g[x\mapsto d]} \phi$*.*

The existential is dual: $\mathfrak{M}, w \Vdash_g \exists x\, \phi$ *iff* $\mathfrak{M}, w \Vdash_{g[x\mapsto d]} \phi$ *for some* $d \in D_w$*.*[4]

The quantifier clause is word for word the naïve one, ranging over D_w. What has changed is the atomic clause: predication now carries its own existential requirement through the residence condition, so a term must denote a resident even to be a candidate for satisfying

[3] This is not to say that there could not be an object in a possible world which in fact is a winged horse, but then this makes it a merely contingent object.

[4] Restall's own system also carries a *wide* or *possibilist* pair of quantifiers Π, Σ ranging over the whole global domain D rather than the local D_w, so that $\Sigma x\, \phi$ holds at w when some $d \in D$, resident or not, satisfies ϕ. The narrow actualist quantifiers are then recovered by guarding: $\forall x\, \phi$ as $\Pi x(E!x \to \phi)$ and $\exists x\, \phi$ as $\Sigma x(E!x \wedge \phi)$. The wide pair lets one say that there is, in the widest sense, a thing that could have been a dodo while still denying that any such thing exists, and it validates the wide Barcan formula $\Pi x\, \Box\phi \to \Box\, \Pi x\, \phi$ that the narrow pair refutes. We keep to the narrow quantifiers in the main text and mention the wide pair only for faithfulness to Restall.

a predicate. The effect shows itself at once in the inference that classical logic treats as the very meaning of the universal quantifier.

4.2 The Cost: Universal Instantiation Fails

In classical predicate logic one may always pass from $\forall x\,\phi$ to $\phi[t/x]$: if everything is ϕ, then any particular thing t is ϕ. The contingentist must give this up, and seeing why is the quickest route into the position.

Proposition 5. *Universal instantiation $\forall x\,\phi \to \phi[t/x]$ is not valid on the contingentist semantics.*

Proof. Let $W = \{w, v\}$ with no accessibility, and $D_w = \{a\}$ and $D_v = \{b\}$. Interpret F by $I(F, w) = \{a\}$; the residence condition permits this, since $a \in D_w$. Then $\forall x\,Fx$ holds at w, the only resident a of D_w being F. But Fb fails at w: the residence condition forbids b, a non-resident, from lying in $I(F, w)$ at all, so $\langle b\rangle \notin I(F, w)$. Hence $\forall x\,Fx \to Fb$ is false at w. □

What replaces the classical rule is its free-logic correction: from $\forall x\,\phi$ one may pass to $\phi[t/x]$ only on the added premise that t exists, that is,

$$\forall x\,\phi \wedge E!t \;\to\; \phi[t/x],$$

which *is* valid, since the existence premise forces $[\![t]\!] \in D_w$, putting t within the quantifier's range. Relatedly, the residence condition ensures that an object can only satisfy an atomic predicate P if it exists.[5] That is, for atomic formulas,

$$P(t) \to E!t$$

is valid. This gives negative free logic a unique feature: while the free-logic rule for existential generalization requires an existence premise ($(\phi[t/x] \wedge E!t) \to \exists x\,\phi$), the residence condition guarantees that atomic formulas satisfy this premise automatically. Consequently, classical existential generalization remains perfectly valid for atomic predicates without any added premises:

$$P(t) \to \exists x\,P(x)$$

.

This is the cost Williamson pressed against the contingentist, and it is real: the clean classical rules are gone, replaced by guarded versions that must check existence at every step. The contingentist's reply, which we register without adjudicating, is that the cost is slight and the company respectable. Mathematics is shot through with terms that may fail to denote– $1/0$, the limit of a divergent sequence, the set of all non-self-membered sets– and the free logic that handles them is thoroughly classical and entirely workable. A term that might have denoted but does not, on this view, is no stranger than a term that might

[5] In negative free logic, all atomic properties are false of non-existent objects. If b does not exist, Fb is false. Because Fb is false, $\neg Fb$ is true. Therefore, a non-existent object can satisfy the complex formula $\neg Fx$, so that is why we restrict to only atomic predicates.

denote but, as the description stands, does not. The guarded rules are the price of taking non-existence seriously, and the contingentist judges it a fair one.

4.3 The First Payoff: Both Barcan Formulae Fail

We can now return to the Barcan formulae and see them fail again. It is the position's central claim made formal. The countermodels are essentially those of Section 2, now reread through the existence predicate, which is what gives the failures their contingentist meaning.

Recall the shape of each failure. The Barcan formula $\forall x \Box\phi \rightarrow \Box\forall x\, \phi$ fails when a world v accessible from w contains an individual not in D_w: the antecedent speaks only of what exists at w, the consequent's inner quantifier reaches the newcomers at v, and a newcomer may fail ϕ. In the contingentist's vocabulary this is the failure of the inference from *everything that actually exists is necessarily* ϕ to *necessarily everything is* ϕ, and it fails precisely because the merely possible objects, the ones that exist at v but not at w, are not bound by a premise about the actual ones. The converse formula fails in the mirror image, when an individual of D_w is absent from some accessible D_v: the contingentist reads this as honouring the thought that what actually exists might not have, so that a thing here need not be governed by a generalisation holding only where it is present.

The dual form sharpens the point and is the one Restall foregrounds. Consider

$$\Diamond \exists x\, Fx \;\rightarrow\; \exists x \Diamond Fx.$$

The antecedent says there could have been an F; the consequent says there actually exists something that could have been an F. The necessitist accepts the inference, because for the necessitist the possible F, existing in the circumstances where it is F, must exist here too. The contingentist denies it, and the countermodel says exactly why.

Proposition 6. *$\Diamond \exists x\, Fx \rightarrow \exists x \Diamond Fx$ is not valid on the contingentist semantics.*

Proof. Let $W = \{w, v\}$ with wRv, $D_w = \{a\}$ and $D_v = \{a, d\}$, where $d \notin D_w$. Interpret F by $I(F, w) = \varnothing$ and $I(F, v) = \{d\}$. For the antecedent: at v the resident d is F, so $\exists x\, Fx$ holds at v, and since wRv, $\Diamond \exists x\, Fx$ holds at w. For the consequent: we need some resident of D_w that is possibly F. The only resident is a, and a's sole accessible world is v, where $a \notin I(F, v) = \{d\}$. So a is not possibly F, and as a is the only candidate, $\exists x \Diamond Fx$ fails at w. The implication is false at w. □

The witness for the possibility was d, an object that exists at v but not at w (see Figure 6). The necessitist would have d already present at w, as a non-concrete existent, available for the outer quantifier; the contingentist refuses to put it there, and the existence predicate is what enforces the refusal, for the outer $\exists$ ranges over D_w alone and $d \notin D_w$. Nor can the negative free logic smuggle d back in through predication: since d is absent from w, the residence condition bars it from every extension at w, so there is not even an atomic truth about d here to trade on. The merely possible F, Restall's possible dodo or possible daughter, is genuinely merely possible: a thing that could have existed and been F, of which nothing whatever is true at a world where it does not exist, and which therefore cannot witness an actual existential. The failure of the Barcan formula is the formal shadow of that conviction.



Figure 6: The contingentist countermodel to $\Diamond \exists x\, Fx \to \exists x \Diamond Fx$. There could have been an F, witnessed by d at v. But d does not exist at w, and the outer quantifier ranges only over $D_w = \{a\}$, whose sole member is not F.

4.4 The Second Payoff: Completeness Recovered†

The Barcan failures were the payoff the contingentist wanted; completeness is the payoff the contingentist needed. Recall that the naïve chapter did not founder on the Barcan formulae alone. It foundered, more quietly and more seriously, on a mismatch between the logic and any calculus for it: for the symmetric systems we care about most, and above all for **QS5**, the canonical construction could not be made to deliver a model of the right shape. It is worth seeing that the free logic just built repairs this too, and that it repairs it in a way exactly dual to the necessitist's.

Return to the point where the naïve construction jammed. We had a world Γ that, containing $\exists x\, Px$ and being maximal, was forced to *name* a witness: some term c with $Pc \in \Gamma$, drawn from what Γ acknowledged to exist. An accessed world Δ then declared $\Box Pd$ for a term d of its own, and symmetry demanded that this claim (Pd) hold at Γ, requiring $Pd \in \Gamma$ for a d that Γ had never certified. The witnesses fixed on the two sides did not coincide, and nothing in the naïve apparatus could make them, because the syntax bound each world's witnesses to that world's domain and left the domains free to disagree.

The free logic dissolves the obstruction at its root by breaking the very link between *naming a term* and *admitting it to the domain* that did the damage. In the naïve setting, for a term c to occur in a world's theory was already for c to count as a resident there; the guarded instantiation rule of Section 4.2 severs exactly this tie. Because passage from $\forall x\, \phi$ to $\phi[c/x]$ now requires the extra premise $E!c$, a maximal consistent theory may mention c, and reason with it freely, without thereby containing $E!c$. This lets us define the canonical domain of a world by existence rather than by mere occurrence:

$$D_\Gamma \;:=\; \{\, c \mid E!c \in \Gamma \,\},$$

while every term, resident or not, takes its value in the single global domain D built from all the terms of the language. The distinction the object language draws with $E!$, between denoting and existing-here, is precisely the distinction the construction now draws between the global domain and the local one. And the residence condition of Definition 9 is honoured by construction: we place a tuple of terms in the canonical extension $I(P, \Gamma)$ only when P of those terms belongs to Γ *and* each term is certified by $E!$ in Γ, so that a canonical predication never reaches outside the local domain.

Now let us watch the backward demand that broke symmetry in the relational canonical model of the Section 2.4, and see why, once predication answers to existence, it no longer

overreaches.

Suppose Δ contains $\Box Pd$, and symmetry requires that Pd must lie in Γ. Here the negative free logic does the decisive work, and it does it *not* by relaxing predication but by controlling what a backward necessity can be about. A necessity $\Box Pd$ that is to be discharged at Γ is discharged there as the plain predication Pd; and for Pd to belong to a maximal consistent Γ that respects the guarded rules, the term d must be one that Γ certifies to exist, that is, $E!d \in \Gamma$, so that $d \in D_\Gamma$. The residence condition, far from obstructing the construction, is what *drives* it: a backward-travelling necessity can only ever concern a resident of the world it returns to. The would-be mismatch of the naïve case cannot arise, since the only necessities that come home to Γ are ones about Γ's own residents.

Two residents, Γ's own witness and the individual the returning necessity concerns, may of course be introduced under different terms; here the construction leans, as the necessitist's did, on the assignment under the interpretaition function. Terms contribute to satisfaction only through what they denote, so two terms with the same global value are interchangeable in every formula.[6] A necessity arriving at Γ under a term d is thus honoured whenever Γ houses the individual $[\![d]\!]$ under any term at all, and the residence condition guarantees that, the necessity being about a resident, Γ does. The backward link that symmetry demands is satisfied by the fact that a negative-free theory only ever sends home necessities its destination can receive.

The moral is worth stating beside the necessitist's, for the two are mirror images down to the mechanism. The naïve construction failed because a witness fixed at one world could fall outside the domain of another, and the logic *cared*. The necessitist removed the failure by removing the difference: one constant domain, so no witness is ever outside it, and the identity apparatus lets a returning individual be recognised as one already present. The contingentist removes the same failure in different way: the domains still vary, and a witness fixed at one world may indeed be absent from another, but the guarded quantifier rules and the residence condition together ensure that no world is ever asked to honour a predication about an object it does not house. The necessitist makes every witness present everywhere; the contingentist makes sure nothing absent is ever demanded.[7]

4.5 A Variation and a Scruple

Positive free logic. We built a *negative* free logic, on which an atomic predication is false unless every term denotes a resident. The natural alternative is a *positive* free logic, which drops the residence condition and lets $I(P, w)$ contain tuples of non-residents, so that a predicate may hold at w of an object absent from D_w. On that choice *Pegasus is a winged horse* comes out true, and one may speak truly, in bare predication, of merely possible objects. The Barcan failures survive either way, as the countermodels of Section 4.3 confirm, turning as they do on the domains and not on the treatment of the absent; what the choice governs is only the verdict on predications about non-residents. We have

[6] This is Restall's *semantic value lemma*: if $[\![t]\!] = [\![s]\!]$ then A and $A[s/t]$ hold together at every world under every assignment. It is the negative-free-logic counterpart of the identity-recognition step by which the necessitist's returning individual was matched to a resident already present.

[7] The construction sketched here is the model-theorist's repair: it builds the accessibility relation relationally, as in the naïve and necessitist chapters, and turns on showing that the backward necessities symmetry demands only ever concern residents of the world they return to. Restall's own completeness proof takes a different route, see the Appendix of (Restall, 2026).

followed Restall in taking the negative road, on which predication answers to existence.

The necessity of identity. Identity behaves delicately here, as we anticipated in Section 2. Where both terms denote residents, the necessity of identity holds as before. But an identity statement about an object absent from the world of evaluation is governed by the residence condition like any other predication: on the negative logic we have adopted, $a = a$ is *false* at a world where a does not exist, for identity is a predicate and a is not there to satisfy it. The contingentist thus says that identities are necessary *where their objects exist*, and makes no commitment to their holding where the objects are absent, which is exactly the flexibility the necessitist's constant domain denied itself. Necessary self-identity becomes the guarded $E!a \to a = a$, true everywhere, rather than the unguarded $a = a$, which fails wherever a is missing.

Exercises

Exercise 22

Warm up on the existence predicate. Work in a contingentist model with $W = \{w, v\}$, wRv, $D_w = \{a\}$, $D_v = \{a, d\}$, and $I(F, w) = \{a\}$, $I(F, v) = \{a, d\}$. (Check first that these extensions obey the residence condition: no tuple contains a non-resident.) Evaluate each formula at w, naming the clause that forces the value, and paying attention to predications and quantifications involving d, which does not exist at w.

(a) $E!a$ and $E!d$

(b) Fd (note the residence condition at work: d is F at v, yet Fd is *false* at w)

(c) $\exists x\, Fx$

(d) $\Diamond\, E!d$

(e) $\forall x\, Fx \to Fd$

Exercise 23

Construct a countermodel, drawn as a labelled diagram, for each of the following, and in each case say in one sentence whether the failure turns on the domain *growing* or *shrinking* along the accessibility relation.

(a) $\forall x \Box Fx \to \Box \forall x\, Fx$ (BF)

(b) $\Box \forall x\, Fx \to \forall x \Box Fx$ (CBF)

(c) $\exists x \Diamond \neg Fx \to \Diamond \exists x \neg Fx$

For (c), explain why this implication, which held under the necessitist's constant domains, now fails: identify the single feature of the contingentist semantics responsible.

Exercise 24

The text proved the dual Barcan formula $\Diamond \exists x\, Fx \to \exists x \Diamond Fx$ invalid (Proposition 6). Complete the family.

(a) State the dual of the converse Barcan formula in the $\Diamond/\exists$ idiom, and give a countermodel.

(b) Verify that your countermodel for (a) is, structurally, the mirror image of the one in the text: where that one used a world that *gains* an individual, yours should use one that *loses* one. Make the correspondence explicit.

Exercise 25

Universal instantiation is the price of free logic. Work through its repair.

(a) Give a one-world countermodel to $\forall x\, Fx \to Fb$, where b is a constant denoting a non-resident of the world's domain.

(b) Prove semantically that the *guarded* version $\big(\forall x\, Fx \wedge E!b\big) \to Fb$ is valid on every contingentist model. Identify the exact step where the existence premise does its work.

(c) State and prove the dual guarded rule for the existential, the corrected form of the inference from Fb to $\exists x\, Fx$.

Exercise 26*

Falsify the necessitist thesis. Recall it may be written

$$\Box \forall x \Box \exists y\, (x = y), \tag{NT}$$

"necessarily, everything is necessarily identical to something".

(a) Construct a contingentist countermodel to (NT). (Hint: you need a world w with a resident a and an accessible world v from which a is absent; the inner $\exists y\, (x = y)$, evaluated at v with x assigned a, then has no resident witness.)

(b) Identify precisely which feature of the necessitist semantics, restored here, blocked this countermodel there. Phrase your answer in terms of the existence predicate.

(c) In one or two sentences, state what the validity of (NT) meant for the necessitist and what its *failure* means for the contingentist, so that the same formula reads as a metaphysical truth for one and a metaphysical falsehood for the other.

Exercise 27

The wide quantifiers of the footnote to Definition 10. Let Π, Σ range over the global domain $D = \bigcup_w D_w$, so $\Sigma x\, \phi$ holds at w when some $d \in D$ satisfies ϕ, resident or not.

(a) Show that $\forall x\, \phi$ and $\Pi x(E!x \to \phi)$ are equivalent on every contingentist model, and likewise $\exists x\, \phi$ and $\Sigma x(E!x \wedge \phi)$, so that the narrow quantifiers are definable from the wide.

(b) Using the wide existential, write a formula saying "there could have been a thing that is F, though no such thing actually exists," and verify it is satisfiable. (Contrast this with the bare Fd of Exercise 1(b): the truth about the merely possible is carried by a modal operator, never by a bare predication.)

(c) Show the wide Barcan formula $\Pi x\,\Box\phi \to \Box\,\Pi x\,\phi$ *is* valid, and say in a sentence why the wide quantifier validates what the narrow one refutes. Does this concede anything to necessitism? Answer briefly.

Exercise 28*

Positive versus negative free logic. The main text adopted a *negative* free logic, imposing the residence condition so that a predication about a non-resident is false. The *positive* alternative drops the condition, letting $I(P, w)$ contain tuples of non-residents.

(a) State the positive-free-logic truth clause for atomic formulas, and say how it differs from the residence-guarded clause of Definition 10.

(b) Exhibit a single model and formula coming out *true* under the positive clause and *false* under the negative one. (*Pegasus is a winged horse* is the intended shape of the example.)

(c) Show that the failure of the Barcan formula survives in *either* free logic, by checking that your BF countermodel from Exercise 2(a) refutes it under both clauses. Conclude that the Barcan failures are a feature of the varying domains, not of the choice between positive and negative free logic.

Exercise 29

Serious actualism is the thesis that nothing can have a property without existing.

(a) Explain why the *negative* free logic of the main text is the natural home of serious actualism, and why the *positive* variant of the previous exercise is in tension with it. (Your answer should turn on the residence condition and what it forbids.)

(b) Consider *Pegasus does not exist*, which the contingentist wants true. Show that this sentence is *not* itself a violation of serious actualism, even though it appears to predicate something of Pegasus. (Hint: distinguish $\neg E!$ Pegasus from an atomic predication; the serious actualist denies that non-existence is a property, not that the sentence is true.)

Exercise 30**

The two responses as mirror images. This exercise asks you to make the symmetry between the sections precise.

(a) Fill in a table with rows for: the treatment of domains, the status of the Barcan formulae, the underlying logic of the quantifiers, the metaphysical thesis incurred, and the consequences. Give the necessitist and contingentist entry for each row.

(b) The necessitist recovers a clean completeness result; the contingentist's logic refuses the Barcan formulae as theorems. Is there a sense in which the contingentist, too, has a well-behaved logic, or has contingentism bought its metaphysics at a genuine logical loss? Argue briefly, drawing on the failure of universal instantiation and its guarded repair.

References

Barcan, R. C. (1946). A functional calculus of first order based on strict implication. *The Journal of Symbolic Logic*, *11*(1), 1–16. https://doi.org/10.2307/2269159

Barcan, R. C. (1947). The identity of individuals in a strict functional calculus of second order. *The Journal of Symbolic Logic*, *12*(1), 12–15. https://doi.org/10.2307/2267171

Gabbay, D. M., Skvortsov, D., & Shehtman, V. (2009). *Quantification in Nonclassical Logic*. Elsevier Science.

Kripke, S. A. (1980). *Naming and Necessity*. Harvard University Press.

Quine, W. V. (1953). Reference and Modality. In *From a Logical Point of View* (pp. 139–159). Harvard University Press.

Restall, G. (2019). Generality and existence 1: Quantification and free logic. *The Review of Symbolic Logic*, *12*(1), 1–29. https://doi.org/10.1017/S175502031800031X

Restall, G. (2026). Modal Logic and Contingent Existence (Generality and Existence 2) [forthcoming]. *Australasian Philosophical Review*.

Searle, J. R. (1958). Proper Names. *Mind*, *67*(266), 166–173. https://doi.org/10.2307/2184031

Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford University Press.