Taming Counterpart Semantics

Wolfgang Schwarz

Incomplete draft, 23 July 2023

Contents

1 Introduction

We turn now to what is arguably one of the least well behaved
modal languages ever proposed: first-order modal logic.
[Blackburn and van Benthem 2007]

1.1 Beyond Kripke semantics

Modal logic has outgrown its philosophical origins. What used to be the logic of possibility and necessity has become topic-neutral, with applications ranging from the validation of computer programs to the study of mathematical proofs.

Along the way, modal predicate logic has lost its role as the centre of investigation, to the point that it is hardly mentioned in many textbooks. Indeed, propositional modal logic has emerged as a fragment of first-order predicate logic, with the domain of "worlds" playing the role of individuals. The distinctive character of modal logic – emphasized in [Blackburn et al. 2001] – is not its subject matter, but its perspective. Statements of modal logic describe relational structures from the inside perspective of a particular node. Modal predicate logic emerges as a somewhat cumbersome hybrid, combining an internal perspective on one class of objects (the domain of the modal operators) with an external perspective on a possibly different class of objects (the domain of quantification).

Cumbersome though it may be, this hybrid perspective is useful and natural for many applications. When reasoning about time, for example, it makes sense to take a perspective that is internal to the structure of times (so that what is true at one point may be false at another), but external to the structure of sticks and stones and people that exist at the various times.

According to standard "Kripke semantics" for modal predicate logic, modal operators shift the world of evaluation, and thereby the extension of predicates, but they

don't affect the reference of singular terms. $\exists x \Diamond Fx$ is true at a world w iff there is an individual (at w) that satisfies Fx at some world accessible from w'.

While conceptually simple and possibly adequate for some applications, there are reasons to look for alternatives.

One well-known alternative is *individual concept semantics*, in which singular terms don't merely pick out individuals, but express functions from world to individuals. This essay focusses on an alternative that retains the purely referential interpretation of singular terms but allows modal operators to shift their reference: In *counterpart semantics*, $\exists x \Diamond Fx$ is true at a world w iff the there is an individual (at w) one of whose *counterparts* at some accessible point w' satisfies Fx.

David Lewis introduced counterpart semantics (in [Lewis 1968]) mainly because his "realist" interpretation of possible worlds did not allow an individual to be part of multiple worlds. Even philosophers who reject Lewis's realism have come to appreciate counterpart semantics for its power to solve philosophical puzzles about identity across worlds and times. (See, for example, [Lewis 1971], [Hawley 2001], [Divers 2007], [Schwarz 2014], [McDonnell 2016], [Ninan 2018], [Kocurek 2018], [Ramachandran 2020], [Hicks 2022].)

In the meantime, mathematical logicians have found other reasons to study alternatives to Kripke semantics. While Kripke semantics for *propositional* modal logic has established itself as a potent framework to study a wide range of logical systems, the situation in modal predicate logic is comparatively bleak. Establishing completeness requires different techniques for different types of systems (as reviewed, for example, in [Hughes and Cresswell 1996: chs.14-17] and [Garson 2001: ch.2]), and many important systems remain incomplete (see [Ono 1973] [Ghilardi 1991], [Cresswell 1995], [Gasquet 1995], [Belardinelli 2022]).

Notable alternatives to Kripke semantics developed by mathematical logicians include the functor semantics of [Ghilardi 1989], the metaframe semantics of [Skvortsov and Shehtman 1993], and the hyperdoctrinal semantics of [Shirasu 1998]. Without the simplicity and intuitive appeal of Kripke semantics, however, these have gained no traction in philosophy.

All three alternatives bear a certain abstract resemblance to counterpart semantics This suggests that some form of counterpart semantics might provide an alternative to Kripke semantics that simultaneously addresses the latter's technical and philosophical shortcomings.

I say 'some form of' counterpart semantics because the original semantics of

[Lewis 1968] arguably won't do. The logic determined by Lewis's semantics is strange – too weak in some respects and too strong and in others. It is too strong insofar as it validates controversial principles such as the Necessity of Existence and the Converse Barcan Formula. It is too weak insofar as it fails to validate basic principles of propositional modal logic such as $\Box(A \wedge B) \supset \Box A$. It also requires mysterious restrictions to classical substitution principles: $\forall x \Diamond Gxy$, for example, does not entail $\Diamond Gyy$. (See [Hazen 1979] and [Woollaston 1994], among others.)

There are two main sources of this deviance.

The first is Lewis's choice to interpret the box as "strong necessity". In effect, he thereby swapped the traditional, hybrid perspective of Kripke semantics (looking at individuals from the outside but at the structure of worlds from the inside) for a thoroughly internal perspective that takes modal contexts to express properties of individuals, rather than the worlds they inhabit (compare [Lewis 1986: 230-235]). As I'll explain in the next section, this calls for a change to the syntax of modal predicate logic. Silvio Ghilardi and Giancarlo Meloni have used a "typed" language (as advertised by [Lawvere 1969]) in their formalisation of counterpart semantics (see, e.g., [Ghilardi and Meloni 1988], [Ghilardi and Meloni 1991], and [Ghilardi 2001]). Good introductions to this approach are [Corsi 2002a], [Belardinelli 2006], and [Braüner and Ghilardi 2007: 591–616].

The other source of deviance for Lewis's logic is that he allows individuals to have multiple counterparts at some accessible world. This, too, arguably calls for a change – or at least an extension – to the syntax of modal predicate logic. In Ghildardi's approach, modal formulas function syntactically like predicates. $\Diamond Gxy$, for example, is a binary predicate that combines with two terms u, v to the formula $(\Diamond Gxy)uv$. The rule of universal instantiation only allows replacing variables outside the modal predicate: from $\forall x(\Diamond Gxy)xv$ one can infer $(\Diamond Gxy)uv$, but not $(\Diamond Guy)uv$.

All this works out fine. But what if we want to stick to the standard syntax of modal predicate logic, and to the hybrid perspective on modal structures? Can we still use counterpart semantics, without buying into all the deviance of Lewis's proposal?

This question was raised by Allen Hazen in [Hazen 1977] and [Hazen 1979]. Hazen showed that one can indeed tweak Lewis's semantics to obtain the familiar logic of Kripke semantics. More recently, [Kutz 2000] and [Kracht and Kutz 2002] have offered a more straightforward counterpart semantics for the standard language of modal predicate logic, building on the metaframe semantics of [Skvortsov and Shehtman 1993]. Unlike Hazen, their aim is not to recover the same logic as Kripke

semantics, but to obtain a more versatile model theory that can be applied to a wider range of logical systems. (See also [Kracht and Kutz 2005] and [Kracht and Kutz 2007].)

Unfortunately, some key claims in [Kutz 2000] and [Kracht and Kutz 2002] are incorrect. The logic described by Kutz and Kracht is not complete with respect to their semantics, and their approach to completeness does not work for many important systems.

In this essay, I will adopt and extend the "straightforward" approach chosen by Kutz and Kracht. I will describe the minimal logics determined by different variants of this account, and explain how imposing restrictions on the counterpart relations can determine stronger logics – including some that are incomplete in Kripke semantics.

The main point I want to establish is that counterpart semantics can be tamed. The logic of [Lewis 1968] is deviant, but its deviance is not an inevitable aspect of counterpart semantics. On the contrary, by relaxing Kripke's assumption of strict transworld identity, one can obtain an intuitive and philosophically attractive model theory for many important systems of modal predicate logic.

1.2 Overview and Apology

I am sorry about the length of this essay. I'm not a logician, and it is easy for me to make mistakes in this territory. I have therefore spelled out many proofs in rather tedious detail, and I have not used shortcuts a more knowledgeable author might have found.

Let me give an overview of what is to come.

Chapter 2 introduces counterpart models as generalised Kripke models in which the relation of strict transworld identity is replaced by a counterpart relation – or rather, by a family of counterpart relations.

Kripke semantics comes in many flavours, and this diversity is inherited by counterpart semantics. One important choice point (in Kripke semantics) is how to deal with individuals that go out of existence as the point of evaluation moves from world to world. One option is to stipulate that this never happens, leading to *constant domain* or *expanding domain semantics*. If one doesn't want to make this stipulation, the underlying first-order logic should arguably be weakened to a free logic. Dif-

ferent types of free logic recommend themselves depending on whether individuals that don't exist at a world can differ in which predicates they satisfy at this world. A *positive* answer requires associating each world with two domains, an "inner" domain of existing individuals and an "outer" domain of nonexistent individuals. A *negative* answer does not require a separate outer domain; individuals are assumed not to satisfy any atomic predicates at worlds where they don't exist.

The same choices arise in counterpart semantics, except that what matters here is not so much whether an individual itself can be found in the domain of accessible worlds, but whether it has a counterpart there. I am going to explore all three options: a "classical" semantics in which individuals have counterparts at all accessible worlds, a "positive" semantics with inner and outer domains, and a "negative" semantics in which nonexistent individuals don't satisfy atomic predicates.

In chapters 3–6, I will impose a further condition on counterparts: I will focus on *functional* models in which a single individual never has multiple counterparts at the same world, except relative to different counterpart relations. This allows us to retain the traditional substitution principles of first-order logic. If we add the assumption that individuals never go out of existence – meaning that every individual has at least one counterpart at every accessible world – the resulting logic is a simple combination of classical first-order logic with the minimal modal logic K. If we allow individuals to go out of existence and adopt a positive approach, we get a combination of standard positive free logic with K. On the negative approach, the logic is an extension of standard negative free logic combined with K. Soundness is proved in chapter 3, completeness in chapter 4.

Chapter 4 also explains how counterpart semantics gets around an obstacle to the construction of canonical models in Kripke semantics. In Kripke semantics, the central "truth lemma" requires different constructions of canonical models for different logics. By contrast, the simple construction in chapter 4 allows proving the truth lemma for any extension of our base logics.

Unfortunately, many interesting systems are not sound on the structure of their canonical model, given the construction from chapter 4. Chapter 6 begins to explore alternative constructions that may be used to prove completeness for some of these systems.

In between, in chapter 5, I briefly discuss the "correspondence" between modal formulas (or schemas) and properties of counterpart structures. In Kripke semantics, the (T) schema $\Box A \supset A$, for example, corresponds to reflexivity of the accessibility

relation, in the sense that the schema is valid on a frame iff the relation is reflexive. In counterpart semantics, the schema instead defines a joint property of the accessibility relation and the counterpart relation(s).

In chapters 7–9, I drop the functionality assumption. I now allow individuals to have multiple counterparts at the same world. It is well-known that this breaks the substitution principles of classical (and free) logic, but it is not well-known exactly how these principles break, and under what restrictions they still hold. This question is answered in chapter 7, although completeness is only proved in chapter 9. The proof requires a more complicated construction of canonical models than the one used in chapter 4.

The failure of traditional substitution principles is a sign that the standard language of modal predicate logic lacks the expressive resources to adequately talk about non-functional structures. The missing power could be restored by adding lambda abstraction. In chapter 8, I explore an option that addresses the problem more directly. I here add an object-language substitution operator to the language and describe its logic. The completeness proofs in chapter 9 cover both systems in the original language and systems in the extended language.

> Much of this essay was written in 2009–2010, with major revisions in 2022–2023. While I hope that the main results are correct, some important questions remain open. Chapter 6 in particular is a mere sketch. I have added red boxes like this one at points where further work is needed.
>
> Two more apologies. I realize that many of the proofs are not only tedious, but inelegant and sometimes hard to follow. I have not (yet) bothered rewriting them once I convinced myself that the result is correct. Also, I should add more references to (and discussions of) the relevant literature.

2 Counterpart Models

2.1 Worlds and individuals

There are two obvious ways of combining the standard model theory of first-order logic with that of propositional modal logic. We can either have a single domain of individuals, or we can associate each world with its own domain, so that different individuals can "exist at" different worlds.

In Kripke semantics, the first option – "constant domain semantics" – renders the Barcan Formula

(BF) $\forall x \Box A \supset \Box \forall x A$

valid, even though (BF) is generally not provable in systems that combine the rules of first-order logic with those of the minimal modal logic K. Constant domain semantics therefore can't provide a model theory for many natural systems of modal predicate logic.

If we allow for variable domains, we should arguably weaken the classical logic of quantification. In classical logic, $\exists x\, x = a$ is a logical truth. Since normal modal logics are closed under necessitation, we could infer $\Box \exists x\, x = a$. By universal generalization, we derive the "Necessity of Existence",

(NE) $\forall y \Box \exists x\, x = y.$

If different individuals can exist at different worlds, however, we probably don't want (NE) to be valid.

Free logics are weakenings of classical logic in which $\exists x\, x = a$ is not provable: singular terms are not required to refer to (existing) individuals. The question now arises if anything interesting can be said about nonexistent individuals. Could we have $Fa \land \neg Fb$, even though 'a' and 'b' are both empty? *Positive* free logic allows for this. Here we need to distinguish an *inner* domain D of existing individuals –

over which the quantifiers range – and an *outer* domain U of individuals that may not exist but can still serve as referents of singular terms and fall into the extension of atomic predicates. *Negative* and *nonvalent* approaches, by contrast, assume that atomic predications with empty terms are always either false or indeterminate. Here we don't need an extra outer domain.

Each of these options can be used in the semantics of modal predicate logic. In this context, the outer domain of positive free logic is naturally understood as the domain of "possible individuals" – the union of the (inner) domains of all worlds. (See, for example, [Kripke 1963].)

At the opposite end of the spectrum from constant domains to variable domains lies Lewis's [1968, 1986] requirement of *disjoined domains*. According to Lewis, no ordinary individual exists at more than one world. We are not going to adopt this assumption. Counterpart semantics, as here developed here, is compatible with constant domains, disjoined domains, and merely variable domains.

As in Kripke semantics, we will have slightly different models depending on whether the underlying theory of quantification is classical or free, and on whether the free logic is positive or negative. (Extension to nonvalent approaches is straightforward, but we will not pursue it.) The most general models for a positive approach associate each world w with an inner domain D_w and an outer domain U_w. In models for classical and negatively free logic, we only need the single domains D_w.

Whether these domains are constant or variable (or disjoined) is not terribly important, as we are not going to track individuals by identity. What's more important is how the individuals in the various domains are linked by the counterpart relation. Or, rather, by the counterpart relations – plural.

2.2 Counterpart relations

Lewis introduced counterpart relations as two-place relations between world-bound individuals. Since we allow the same individual to exist at more than one world, we generalise this to a four-place relation between individuals at worlds, so that d_1 at w_1 can be a counterpart of d_2 at w_2 but not of d_2 at w_3. Instead of having a single four-place relation, however, it proves convenient to associate each pair of worlds w, w' with a two-place counterpart relation $C_{w,w'}$, telling us which things in w' are counterparts of which things in w.

In fact, we are going associate each pair of worlds with a *set $K_{w,w'}$* of such relations, in order to account for what Allen Hazen calls "internal relations" (see [Hazen 1979: 328–330], [Lewis 1986: 232f.]).

Suppose Dee and Dum are siblings. Imagine a world w that embeds two copies of the actual world, a "left" copy and a "right" copy. We may want to say that w contains two counterparts of Dee and two of Dum, and that Dee and Dum are necessarily siblings, even though not all counterparts of Dee and Dum at w are siblings of one another.

To model this sort of situations, we need to distinguish different ways of locating individuals from one world in another world. In the example, we can do this by having two counterpart relations, one linking Dee and Dum with their counterparts in the left copy, the other with their counterparts in the right copy. $\Box Gab$ will be true iff, *relative to every counterpart relation*, all counterparts of a are G-related to all counterparts of b.

Having multiple counterpart relations makes no difference to the base logic characterized by our semantics, but it will be useful in the construction of canonical models for stronger logics.

It may help to think of the set $K_{w,w'}$ of counterpart relations as determining a single counterpart relation between sequences of individuals:

> A $\langle d_1', \ldots, d_n' \rangle$ at w' is a counterpart of $\langle d_1, \ldots, d_n \rangle$ at w iff there is a $C \in K_{w,w'}$ such that $d_1 C d_1', \ldots, d_n C d_n'$.

$\langle \text{Dee}_L, \text{Dum}_L \rangle$ and $\langle \text{Dee}_R, \text{Dum}_R \rangle$, for example, may be counterparts of $\langle \text{Dee}, \text{Dum} \rangle$, but $\langle \text{Dee}_L, \text{Dum}_R \rangle$ is not.

That counterparthood should be extended to sequences is suggested in [Lewis 1983] and [Lewis 1986], in response to the problem of internal relations. We could have taken a single counterpart relation between sequences as primitive, but we would then have had to impose some conditions on this relation. For example, we'd want to rule out that a pair $\langle d_1, d_2 \rangle$ at w has $\langle d_1', d_2' \rangle$ at w' as counterpart even though d_1 at w does not have d_1' at w' as counterpart. We'd also want to ensure that if $\langle d_1', d_2' \rangle$ at w' is a counterpart of $\langle d_1, d_2 \rangle$ at w then $\langle d_2', d_1' \rangle$ at w' is a counterpart of $\langle d_2, d_1 \rangle$ at w. These (and other) restrictions are automatically satisfied if we take a set $K_{w,w'}$ of individual counterpart relations as primitive, and derive the counterpart relation between sequences in the way just described.

Now recall that we have different kinds of models, depending on the underlying logic of quantification. If the logic is *classical*, the Necessity of Existence is a theorem. Each world w is associated with a (single) domain D_w of individuals, and we stipulate that if w has access to w' then every member of D_w has a counterpart in $D_{w'}$, relative to every counterpart relation in $K_{w,w'}$. (Otherwise $\forall x \Box \exists y\, x = y$ would be false at w.)

For *positive free logic*, we add an outer domain U_w of individuals to each world. Informally, these are individuals that can be talked about at w, even though they don't exist there. We stipulate that if w has access to w' then every member of U_w has a counterpart in $U_{w'}$. For *negative free logic*, we don't have extra outer domains and we allow that an individual in D_w has no counterpart in $D_{w'}$. (Note that models for classical logic are a special case of the other two kinds of models.)

In the next few chapters, we will impose another restriction on $K_{w,w'}$. We are going to assume that each counterpart relation $C \in K_{w,w'}$ is a possibly partial *function*, so that an individual at one world never has multiple counterparts at another world relative to the same counterpart relation. We still allow for multiple counterparts relative to different counterpart relations, as in the case of Dee and Dum. The functionality assumption ensures that the logic determined by our semantics satisfies standard substitution principles of first-order logic. This will be explained in chapter 7, where the assumption will be lifted.

2.3 Weak necessity and strong necessity

Lewis [1968] gave his semantics in the form of translation rules from the language of modal predicate logic into a non-modal first-order language with explicit quantifiers over worlds and individuals. $\Box Fx$, for example, is translated into a first-order formula stating that every counterpart of x at every world is F. Since every individual at any worlds is identical to itself, $\forall y \Box \exists x\, x = y$ is translated into a logical truth: (NE) comes out as valid.

By letting the box in $\Box \phi(x)$ quantify over counterparts of x, Lewis adopted what Kripke [1971: 137] called a *weak* reading of necessity. Consider a statement like

>   Aristotle is necessarily human,

understood as a metaphysical hypothesis. On its "weak" reading, the statement asserts (informally speaking) that Aristotle is human at every world *at which he exists*.

On its "strong" reading, the statement asserts that Aristotle is human *at every world whatsoever*. Since every individual exists at every world at which it exists, the weak reading validates the Necessity of Existence.

Much of the deviance of Lewis's logic arises from his weak reading of the box. We can see, for example, why the weak reading does not license the inference from $\Box(Fx \wedge Fy)$ to $\Box Fx$. Informally, the premise $\Box(Fx \wedge Fy)$ states that x and y are F at every world at which they both exist. This can be true even if x is $\neg F$ at certain worlds where x exists but y does not, in which case the conclusion $\Box Fx$ is false. In $\Box(Fx \wedge Fy)$, the box effectively ranges only over worlds where x and y both exist (for Lewis: at which they both have counterparts). In $\Box Fx$, the box ranges over the wider set of worlds where x exists, regardless of whether y exists there as well.

The weak reading of the box not only determines a deviant modal logic. It is also hampered by the standard syntax of modal predicate logic. Consider another example (from [Baldwin 1984: 254], see also [Hunter and Seager 1981]).

> Elizabeth is necessarily the daughter of George, but George isn't necessarily the father of Elizabeth.

As a statement about weak necessity, this is easily understood: given the necessity of origin, Elizabeth could not have existed without being the daughter of George, but George could well have existed without having any offspring. But how could the statement be expressed in the language of modal predicate logic? We need a way of saying that Fxy is *necessary for x*, while the same formula is not *necessary for y*.

The sentence operators of modal predicate logic are ill suited for expressing weak necessity. It would be better to use modal predicate operators, as in [Baldwin 1984], or the indexed operators of [Corsi 2007]. The logical deviance of weak necessity would then become more intelligible. Suppose we read $\Box_{x,y}$ as 'it is necessary for x and y that'. We would then say that $\Box_{x,y}(Fx \wedge Fy)$ entails $\Box_{x,y}Fx$, but not $\Box_x Fx$.

One can achieve a similar effect by adding indices to entire sentences, as in the categorical approach of [Lawvere 1969]. Here we would write

$$\Box(Fx \wedge Fy) : x, y$$

instead of $\Box(Fx \wedge Fy)$. The index 'x, y' would make clear that the statement is about x and y. From $\Box(Fx \wedge Fy) : x, y$ we could infer $\Box Fx : x, y$, but not $\Box Fx : x$. [Corsi 2002a], [Belardinelli 2006] and [Braüner and Ghilardi 2007] develop counterpart

semantics for this kind of indexed language, drawing on [Ghilardi and Meloni 1988] and [Ghilardi and Meloni 1991].

It would be wrong, however, to think that counterpart semantics requires a weak reading of the box, just as it would be wrong to think that it requires disjoined domains. We are going to develop a version of counterpart semantics that adopts the strong reading, familiar from Kripke semantics. We don't need an indexed language or indexed operators.

[Lewis 1968] did consider giving the box a strong reading. Admitting that the validity of (NE) may be undesirable, he reports an alternative translation scheme suggested to him by David Kaplan. The alternative translates $\Box \phi(x)$ into 'at every world, some counterpart of x satisfies $\phi(x)$'. Lewis rightly points out that this would have unacceptable consequences for cases in which ϕ is negated: the mere fact that x contingently exists would make $\Diamond Fx$ (i.e., $\neg \Box \neg Fx$) true, for any predicate F.

It is, in fact, hard to provide a semantics for strong necessity with Lewisian translation rules. But the use of translation rules was another of Lewis's idiosyncrasies. As [Hazen 1979] points out, one can convert Lewis's rules into a more standard model-theoretic semantics. In this framework, the strong reading of the box is easily accommodated.

Suppose we evaluate $\Box Fx$ at a world w, relative to an assignment g that maps x to an individual d. In Kripke semantics, $\Box Fx$ is true at w relative to g iff Fx is true at all accessible worlds w', relative to the same assignment g. If d doesn't exist at some such world, we use the resources of free logic to guide our interpretation of Fx. On a positive approach, we assume that d still inhabits the outer domain of w', and that Fx is true at w' depending on whether d is F at w'. On a negative approach, we treat x as genuinely empty at w', and infer that Fx is false at w'.

We have the same options in counterpart semantics. We'll say that $\Box Fx$ is true at w relative to g iff Fx is true at all accessible worlds w' relative to all assignments g' that shift the reference of x to a counterpart of d at w'. If d doesn't have a counterpart at w', we can let $g'(x)$ be undefined. We then stipulate that Fx is false relative to any assignment that leaves x undefined. This is the negative approach. Alternatively, we can adopt a positive approache and stipulate that d must have a counterpart in the outer domain of any accessible world. Then Fx may be true or false at w', depending on whether d's counterparts are at w' are F.

2.4 Models

With all that chit chat out of the way, let's define our models. A *model* combines a *structure* with an interpretation of the language.

Definition 2.1 (Counterpart structure).

A *counterpart structure* consists of

1. a non-empty set W (of "worlds"),
2. a binary ("accessibility") relation R on W,
3. an ("outer domain") function U that assigns to each $w \in W$ a non-empty set U_w,
4. an ("inner domain") function D that assigns to each $w \in W$ a set $D_w \subseteq U_w$, and
5. a ("counterpart-inducing") function K that assigns to each pair of points $\langle w, w' \rangle \in R$ a non-empty set $K_{w,w'}$ of ("counterpart") relations $C \subseteq U_w \times U_{w'}$.

As I've explained in the previous sections, we need the outer domains if we want to say that individuals can differ in their properties even at worlds where they don't exist – meaning, where they don't have counterparts. In this case, we'll assume that all counterpart relations are total, so that an individual always has at least one counterpart at all accessible worlds, if only in the outer domain. If we don't want to distinguish between non-existent individuals, we can use *single-domain* structures without a separate outer domain.

Definition 2.2 (Types of structure).

Let $\mathfrak{S} = \langle W, R, U, D, K \rangle$ be a counterpart structure.

- $\mathfrak{S}$ is *single-domain* if $D_w = U_w$ for all $w \in W$.
- $\mathfrak{S}$ is *total* iff all counterpart relations are total, meaning that any $C \in K_{w,w'}$ relates each $d \in U_w$ to at least one $d' \in U_{w'}$.
- $\mathfrak{S}$ is *functional* iff all counterpart relations are (partial) functions: any $C \in K_{w,w'}$ relates each $d \in U_w$ to at most one $d' \in U_{w'}$.
- $\mathfrak{S}$ is *classical* iff it is single-domain, total, and functional.

The label "classical" alludes to the fact that the logic of classical structures is the

union of classical first-order logic and the minimal normal modal logic K.

Before we can show facts like this, we need to say how sentences of modal predicate logic are interpreted on counterpart structures.

Definition 2.3 (The language of modal predicate logic).

We assume that there is a denumerable list of *predicates*, each associated with an arity, a denumerable list of *(individual) variables*, and a denumerable list of *(individual) constants* (aka *names*). The *formulas* (a.k.a. *sentences*) of $\mathfrak{L}$ are generated by the rule

$$Pt_1 \ldots t_n \mid t_1 = t_2 \mid \neg A \mid (A \supset B) \mid \forall x A \mid \Box A,$$

where P is a predicate with arity n, $t_1, \ldots, t_n$ are *terms* (variables or constants), and x is a variable.

Formulas involving '$\wedge$', '$\vee$', '$\leftrightarrow$', '$\exists$' and '$\Diamond$' are defined by the usual metalinguistic abbreviations. I will often omit parentheses, assuming that the order of precedence among connectives is $\neg, \wedge, \vee, \supset$, with association to the right. I am going to use 'x', 'y', 'z' (sometimes with indices or dashes) as placeholders for arbitrary variables, 'a', 'b', 'c' for names, 't', 's', 'r' for terms, and 'F', 'G', 'P' for predicates with arity 1, 2 and n, respectively. For any expression or set of expressions A, $\mathrm{Var}(A)$ is the set of variables in (members of) A, and $\mathrm{FV}(A)$ is the set of variables with free occurrences in (members of) A.

Definition 2.4 (Interpretation).

Let $\mathfrak{S} = \langle W, R, U, D, K \rangle$ be a counterpart structure. An *interpretation I* (*for $\mathfrak{L}$*) *on* $\mathfrak{S}$ is a function I that assigns to each world $w \in W$ a function I_w such that

 (i) for every non-logical predicate P with arity n, $I_w(P) \subseteq U_w^n$, and
 (ii) $I_w(=) = \{\langle d, d \rangle : d \in U_w\}$.

(For zero-ary predicates P, clause (i) says that $I_w(P) \subseteq U_w^0$. For any U_w, there is exactly one "zero-tuple" in U_w^0, which we may identify with the empty set. So U_w^0 has exactly two subsets, the empty set $\emptyset = 0$ and the unit set of the empty set $\{\emptyset\} = 1$. It is convenient to think of these as truth-values.)

> **Definition 2.5 (Counterpart model).**
> A *counterpart model* $\mathfrak{M}$ consists of a counterpart structure $\mathfrak{S}$ and an interpretation I on $\mathfrak{S}$.

I will call models *single-domain*, *total*, *functional* or *classical* in accordance with their underlying structure.

Formulas are true in a counterpart model relative to a world and an assignment.

> **Definition 2.6 (Assignment).**
> An *assignment on* a set U is a function g from the $\mathfrak{L}$-terms into U.

In negative semantics, we will allow assignments to be partial, otherwise they have to be total.

A few comments on the treatment of names.

According to definition 2.6, assignments interpret not only the variables but also the names of $\mathfrak{L}$. A more conventional treatment would move the interpretation of names out of the assignment and into a model's interpretation function. This could easily be done, but the present treatment proves a little more convenient for our purposes.

For one thing, in a positive semantics, we want each name to pick out an individual in the outer domain of the world at which a formula is evaluated. If models don't contain a designated "actual world" then on which domain U_w should a model interpret a name? We would either (a) have to give names an intensional meaning $W \to U_w$, even though this intension plays no role in the compositional semantics, or (b) include a designated actual world in the structures, or (c) assume that all worlds have the same outer domain.

Another reason for letting names be interpreted by the assignment function is that modal operators shift the interpretation of names just as they shift the interpretation of variables. It is convenient to keep track of shiftable parameters as separate index coordinates, and this way we only need one index coordinate for both names and variables.

In effect, we are treating names as free variables. This has a third advantage. Many classical treatments of modal predicate logic – from [Kripke 1963] and [Lewis 1968] to [Hughes and Cresswell 1996], [Fitting and Mendelsohn 1998], and [Kracht and Kutz 2002] – don't have names in their object language. They do, of course, have

variables. If we think of names as free variables, comparisons with those treatments become easy. (For languages with names, Kracht and Kutz [2005, 2007] switch from counterpart semantics to what Schurz [2011] calls "worldline semantics" – a cousin of individual concept semantics.)

> In an earlier version of this essay, I didn't have a separate category of names. In later chapters, I sometimes still use 'variable' to mean 'singular term'. This should be corrected, but it doesn't affect any results. Throughout this essay, names behave *exactly* like variables that happen to never be bound.

Note also that names and variables are "objectual". They simply pick out individuals. In applications of counterpart semantics, it is sometimes useful to have different counterpart relations for different sorts of individuals, so that one can distinguish, for example, between *person* counterparts and *body* counterparts (see [Lewis 1971]). One may then want to associate each term with a sort, and each sort with its own type of counterpart relation. Here I will focus on the simplest case, where all individuals are of the same sort.

> Many authors have outlined versions of counterpart semantics that allow for different types of counterpart relation. I don't think anyone has worked out the details of the resulting logics. It would be worthwile to do so.

2.5 Truth

Let's spell out what it takes for an $\mathcal{L}$-sentence to be true relative to a world and an assignment in a model.

For the semantics of quantifiers, we need the usual concept of a *variant* of an assignment.

> **Definition 2.7 (Variant).**
> If g is an assignment on some set U and $d \in U$, then the *x-variant* $g^{x \mapsto d}$ is the assignment on U that maps x to d and all other terms t to their original value $g(t)$.

In Kripke semantics, the box shifts the world of evaluation. In counterpart semantics, it also shifts the assignment: $\Box A$ is true relative to w, g iff A is true relative to w', g' for all w', g' such that w' is accessible from w and g' assigns to each term a counterpart at w' of its original value. Let's abbreviate this relationship between w, g and w', g' as $w, g \triangleright w', g'$.

Definition 2.8 (Image).
Let $\mathfrak{S} = \langle W, R, U, D, K \rangle$ be a counterpart structure, w, w' worlds in W, and g, g' assignments on $U_w, U_{w'}$ respectively. We say that g' *at w'* is an *image of g at w* (for short, $w, g \triangleright w', g'$) iff there is a $C \in K_{w,w'}$ such that for every term t, if $g(t)$ is C-related to some $d \in U_{w'}$ then $g(t)Cg'(t)$, otherwise $g'(t)$ is undefined.

In total structures, $g(t)$ is always C-related to some $d \in U_{w'}$, so we have: $w, g \triangleright w', g'$ iff there is a counterpart relation $C \in K_{w,w'}$ that relates $g(t)$ to $g'(t)$ for every term t. We don't need to mention that w' is accessible from w: definition 2.1 ensures that if w' is not accessible from w' then there is no $C \in K_{w,w'}$ at all.

Definition 2.9 (Truth).
Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ be a counterpart model, w a member of W, and g an assignment on U_w. For any predicate P, terms $t_1, \dots, t_n$, and $\mathfrak{L}$-formulas A, B, we define:

$\mathfrak{M}, w, g \Vdash Pt_1 \dots t_n$ iff $\langle g(t_1), \dots, g(t_n) \rangle \in I_w(P)$.

$\mathfrak{M}, w, g \Vdash \neg A$ iff $\mathfrak{M}, w, g \nVdash A$.

$\mathfrak{M}, w, g \Vdash A \supset B$ iff $\mathfrak{M}, w, g \nVdash A$ or $\mathfrak{M}, w, g \Vdash B$.

$\mathfrak{M}, w, g \Vdash \forall x A$ iff $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for all $d \in D_w$.

$\mathfrak{M}, w, g \Vdash \Box A$ iff $\mathfrak{M}, w', g' \Vdash A$ for all w', g' such that $w, g \triangleright w', g'$.

When we're dealing with functional models, the following formulation of the semantics for $\Box A$ is sometimes useful.

Lemma 2.1.
*If $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ is a functional counterpart model, w a member of
W, and g an assignment on U_w, then*

$$\mathfrak{M}, w, g \Vdash \Box A \text{ iff } \mathfrak{M}, w', C \circ g \Vdash A \text{ for all } w', C \text{ s.t. } wRw' \text{ and } C \in K_{w,w'}.$$

Here, $C \circ g$ is the composition of C and g: the function that maps any term t to
$C(g(t))$. (If either $g(t)$ or $C(g(t))$ is undefined, then $(C \circ g)(t)$ is undefined.

Proof. By definition 2.9, $\mathfrak{M}, w, g \Vdash \Box A$ iff $\mathfrak{M}, w', g' \Vdash A$ for all w', g' such that
$w, g \rhd w', g'$. By definition 2.8, $w, g \rhd w', g'$ iff there is a $C \in K_{w,w'}$ such that
for every term t, if $g(t)$ is C-related to some $d \in U_{w'}$ then $g(t) C g'(t)$, otherwise
$g'(t)$ is undefined. Assume C is functional. We can then write $g(t) C g'(t)$ as
$g'(t) = C(g(t)) = (C \circ g)(t)$. So we have $w, g \rhd w', g'$ iff there is a $C \in K_{w,w'}$
such that for every term t,

$$g'(t) = \begin{cases} (C \circ g)(t) & \text{if } \{d : g(x) C d\} \neq \emptyset, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

If $\{d : g(x) C d\} = \emptyset$ then $(C \circ g)(x)$ is undefined. So $w, g \rhd w', g'$ iff there is a
$C \in K_{w,w'}$ such that $g' = C \circ g$. And so $\mathfrak{M}, w, g \Vdash \Box A$ iff $\mathfrak{M}, w', g' \Vdash A$ for all
w', g' such that $g' = C \circ g$ for some $C \in K_{w,w'}$.

In later proofs, I will sometimes refer to the (obvious) fact the truth-value of a
sentence never depends on the value of terms that aren't free in the sentence. Let's
quickly prove this.

Lemma 2.2 (Locality lemma).
*Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ be a counterpart model, $w \in W$, A an $\mathfrak{L}$-formula
and g, g' assignments on U_w such that $g(t) = g'(t)$ for every term t that is
free in A. Then*
$$\mathfrak{M}, w, g \Vdash A \text{ iff } \mathfrak{M}, w, g' \Vdash A.$$

Proof. By induction on (complexity of) A.

- A is atomic. The claim immediately follows from definition 2.9.

- A is $\neg B$. $\mathfrak{M}, w, g \Vdash \neg B$ iff $\mathfrak{M}, w, g \nVdash B$ by definition 2.9, iff $\mathfrak{M}, w, g' \nVdash B$ by induction hypothesis, iff $\mathfrak{M}, w, g' \Vdash \neg B$ by definition 2.9.

- A is $B \supset C$. $\mathfrak{M}, w, g \Vdash B \supset C$ iff $\mathfrak{M}, w, g \nVdash B$ or $\mathfrak{M}, w, g \Vdash C$ by definition 2.9, iff $\mathfrak{M}, w, g' \nVdash B$ or $\mathfrak{M}, w, g' \Vdash C$ by induction hypothesis, iff $\mathfrak{M}, w, g' \Vdash B \supset C$ by definition 2.9.

- A is $\forall x B$. By definition 2.9, $\mathfrak{M}, w, g \Vdash \forall x B$ iff $\mathfrak{M}, w, g^{x \mapsto d} \Vdash B$ for all $d \in D_w$. For each $d \in D_w$, $g^{x \mapsto d}$ and $g'^{x \mapsto d}$ assign the same value to every variable in B. So by induction hypothesis, $\mathfrak{M}, w, g^{x \mapsto d} \Vdash B$ for all $d \in D_w$ iff $\mathfrak{M}, w, g'^{x \mapsto d} \Vdash B$ for all $d \in D_w$, iff $\mathfrak{M}, w, g' \Vdash \forall x B$ by definition 2.9.

- A is $\Box B$. By definition 2.9, $\mathfrak{M}, w, g \Vdash \Box B$ iff $\mathfrak{M}, w', g^* \Vdash B$ for all w', g^* with $w, g \triangleright w', g^*$. Since $g(x) = g'(x)$ for all variables x in B, each w'-image of g at w agrees with some w'-image of g' on all variables in B and vice versa. So by induction hypothesis, $\mathfrak{M}, w', g^* \Vdash B$ for all $\mathfrak{M}, w', g^*$ such that $w, g \triangleright w', g^*$ iff $\mathfrak{M}, w', g'^* \Vdash B$ for all w', g'^* such that $w, g' \triangleright w', g'^*$, iff $\mathfrak{M}, w, g' \Vdash \Box B$ by definition 2.9.

2.6 Positive correlates

In this section, I show that a negative model can be "simulated" by a positive model by adding a "null individual" o to the outer domain U_w of every world w and stipulating that o doesn't satisfy any predicates and isn't denoted by any term.

Definition 2.10 (Positive correlate).
The *positive correlate* $\mathfrak{M}^+$ of a counterpart model $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ is the model $\langle W, R, U^+, D, K^+, I \rangle$ with U^+, K^+ constructed as follows.

Let o be an arbitrary individual (say, the smallest ordinal) not in $\bigcup_w D_w$. For all $w \in W$, let $U_w^+ = U_w \cup \{o\}$.

For all $\langle w, w' \rangle \in R$, let $K^+_{w,w'}$ be the set of relations $C^+ \subseteq U^+_w \times U^+_{w'}$, such that for some $C \in K_{w,w'}$, $C^+ = C \cup \{\langle d, o \rangle : d \in U^+_w$ and there is no $d' \in U_{w'}$ with $dCd'\}$.

The *positive correlate* g^+ of an assignment function g is the function that "completes" g by setting

$$g^+(x) = \begin{cases} g(x) & \text{if } g(x) \text{ is defined} \\ o & \text{otherwise} \end{cases}$$

Lemma 2.3 (Truth-preservation across correlates).
Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ *be any counterpart model and* $\mathfrak{M}^+ = \langle W, R, U^+, D, K^+, I \rangle$ *its positive correlate. For any world $w \in W$, assignment g on U_w, and formula A of* $\mathfrak{L}$,

$$\mathfrak{M}, w, g \Vdash A \text{ iff } \mathfrak{M}^+, w, g^+ \Vdash A,$$

where g^+ is the positive correlate of g.

Proof. By induction on A.

- A is $Px_1 \ldots x_n$. By definition 2.9, $\mathfrak{M}, w, g \Vdash Px_1 \ldots x_n$ iff $\langle g(x_1), \ldots, g(x_n) \rangle \in I_w(P)$. Since $\mathfrak{M}$ and $\mathfrak{M}^+$ have the same interpretation function I, we have to show that $\langle g(x_1), \ldots, g(x_n) \rangle \in I_w(P)$ iff $\langle g^+(x_1), \ldots, g^+(x_n) \rangle \in I_w(P)$. If $g(x_i)$ is defined for all $x_1, \ldots, x_n$ then this follows from the fact that $g^+(x_i) = g(x_i)$. If some $g(x_i)$ is undefined then $\langle g(x_1), \ldots, g(x_n) \rangle$ is undefined and not in $I_w(P)$. We then also have $g^+(x_i) = o$. Since $I_w(P)$ doesn't contains any tuples involving o, we have $\langle g^+(x_1), \ldots, g^+(x_n) \rangle \notin g^+(P)$.

- A is $\neg B$. $\mathfrak{M}, w, g \Vdash \neg B$ iff $\mathfrak{M}, w, g \nVdash B$ by definition 2.9, iff $\mathfrak{M}^+, w, g^+ \nVdash B$ by induction hypothesis, iff $\mathfrak{M}^+, w, g^+ \Vdash \neg B$ by definition 2.9.

- A is $B \supset C$. $\mathfrak{M}, w, g \Vdash B \supset C$ iff $\mathfrak{M}, w, g \nVdash B$ or $\mathfrak{M}, w, g \Vdash C$ by definition 2.9, iff $\mathfrak{M}^+, w, g^+ \nVdash B$ or $\mathfrak{M}^+, w, g^+ \Vdash C$ by induction hypothesis, iff $\mathfrak{M}^+, w, g^+ \Vdash B \supset C$ by definition 2.9.

- *A* is $\forall x B$. By definition 2.9, $\mathfrak{M}, w, g \Vdash \forall x B$ iff $\mathfrak{M}, w, g^{x \mapsto d} \Vdash B$ for all $d \in D_w$. For each $d \in D_w$, $g^{+ x \mapsto d}$ is the positive correlate of $g^{x \mapsto d}$. Thus by induction hypothesis, $\mathfrak{M}, w, g^{x \mapsto d} \Vdash B$ for all $d \in D_w$ iff $\mathfrak{M}, w, g^{+ x \mapsto d} \Vdash B$ for all $d \in D_w$, iff $\mathfrak{M}^+, w, g^+ \Vdash \forall x B$ by definition 2.9.

- *A* is $\Box B$. Assume $\mathfrak{M}, w, g \Vdash \Box B$. By definition 2.9, this means that $\mathfrak{M}, w', g' \Vdash B$ for all w', g' with $w, g \rhd w', g'$. We need to show that $\mathfrak{M}^+, w', g^{+'} \Vdash B$ for all $w', g^{+'}$ with $w, g^+ \rhd w', g^{+'}$. So let $w', g^{+'}$ be such that $w, g^+ \rhd w', g^{+'}$. Since g^+ is total and $\mathfrak{M}^+$ a positive structure, $w, g \rhd w', g^{+'}$ implies that for every variable x there is a $C^+ \in K^+_{w,w'}$ with $g^+(x) C^+ g^{+'}(x)$. Let g' be the assignment on $U_{w'}$ that coincides with $g^{+'}$ except that $g'(x)$ is undefined for every variable x for which $g^{+'}(x) = o$. Let $C = \{\langle d, d' \rangle \in C^+ : d \neq o \text{ and} d' \neq o\}$.

  Now let x be any variable. Assume that there are d, d' with $g(x) = d$ and dCd'. Then $g^+(x) = d$ and dC^+d' and thus $g^{+'}(x) \neq o$, as $\langle d, o \rangle \in C^+$ only if there is no d' with $\langle d, d' \rangle \in C$. So $g'(x) = g^{+'}(x)$, and $g(x) C g'(x)$. On the other hand, assume there are no d, d' with $g(x) = d$ and dCd', either because $g(x)$ is undefined or because $g(x) = d$ and the only d' with $\langle d, d' \rangle \in C^+$ is o. Either way, then $g^{+'}(x) = o$, and so $g'(x)$ is undefined.

  W have shown that for all variables x, if there are d, d' with $g(x) = d$ and dCd' then $g(x) C g'(x)$, otherwise $g'(x)$ is undefined. Since $C \in K_{w,w'}$ by construction of K^+ (in definition 2.10), this means that $w, g \rhd w', g'$. But $g^{+'}$ is the positive correlate of g'. So we've shown that whenever $w, g^+ \rhd w', g^{+'}$ then there is a g' such that $g^{+'}$ is the positive correlate of g' and $w, g \rhd w', g'$. We know that $\mathfrak{M}, w', g' \Vdash B$. So by induction hypothesis, $w', g^{+'} \Vdash B$. That is, $w', g^{+'} \Vdash B$ for each $w', g^{+'}$ with $w, g \rhd w', g'$. By definition 2.9, this means that $w, g^+ \Vdash \Box B$.

  In the other direction, assume $\mathfrak{M}^+, w, g^+ \Vdash \Box B$. That is, $\mathfrak{M}^+, w', g^{+'} \Vdash B$ for each $w', g^{+'}$ with $w, g^+ \rhd w', g^{+'}$. We have to show that $\mathfrak{M}, w', g' \Vdash B$ for all w', g' with $w, g \rhd w', g'$. So let w', g' be such that $w, g \rhd w', g'$. Then there is a $C \in K_{w,w'}$ such that for every variable x, either $g(x) C g'(x)$ or $g(x)$ has no C-counterpart at w' and $g'(x)$ is undefined. Let g'^+ be the positive correlate of g'. Let $C^+ = C \cup \{\langle d, o \rangle : d \in U^+_w \text{ and there is no } d' \in U_{w'} \text{ with}$

dCd' }. By definition 2.10, $C^+ \in K^+_{w,w'}$.

For any variable x, if $g(x)Cg'(x)$ then both $g(x)$ and $g'(x)$ are defined and thus $g^+(x) = g(x)$ and ${g'}^+(x) = g'(x)$, by definition 2.10; moreover, then $g^+(x)C^+{g'}^+(x)$, since $C \subseteq C^+$. On the other hand, if $g(x)$ has no C-counterpart at w', so that $g'(x)$ is undefined, then by construction of C^+ and g^+, $g^+(x)$ (which equals $g(x)$ if $g(x)$ is defined, else o) has o as C^+-counterpart at w', and ${g'}^+(x) = o$; so again $g^+(x)C^+{g'}^+(x)$.

So, for every variable x, there is a $C^+ \in K^+_{w,w'}$ with $g^+(x)C^+{g'}^+(x)$, and so $w, g^+ \rhd w', {g'}^+$.

Now we know that $\mathfrak{M}^+, w', {g^+}' \Vdash B$ for all $w', {g^+}'$ with $w, g \rhd w', {g^+}'$. Hence $\mathfrak{M}^+, w', {g'}^+ \Vdash B$. By induction hypothesis, $\mathfrak{M}, w', g' \Vdash B$. So we've shown that whenever $w, g \rhd w', g'$, then $\mathfrak{M}, w', g' \Vdash B$. By definition 2.9, this means that $\mathfrak{M}, w, g \Vdash \Box B$.

3 The base logics

3.1 FK, QK, and N$^+$K

Until chapter 7, we are now going to limit ourselves to functional counterpart structures, where an individual never has multiple counterparts at the same world relative to the same counterpart relation. Among functional structures, we distinguish between *single-domain* structures, *total structures*, and *classical* structures. Classical structures are both single-domain and total. The aim of this chapter is to describe the logics determined by these three types of structures.

In this context, a *logic* (or *system*) is simply a set of formulas that I will describe by recursive clauses corresponding to the axioms and rules of a "Hilbert-style" calculus. The logic of classical structures simply combines standard axioms and rules of classical first-order logic with those of the minimal normal modal logic K. The logic of total (functional) structures weakens the first-order component to standard positive free logic. For single-domain structures, we'll need an extension of negative free logic.

Standard (non-modal) positive free logic can be defined as the smallest set of formulas L that contains

(Taut) all propositional tautologies

as well as all instances of the axiom schemas

(VQ) $A \supset \forall x A$ provided x is not free in A,

(UD) $\forall x (A \supset B) \supset (\forall x A \supset \forall x B)$,

(FUI) $\forall x A \supset (E!t \supset [t/x]A)$,

($\forall$E!) $\forall x E!x$,

($=$R) $x = x$,

(LL) $s = t \supset (A \supset [s/t]B)$,

and that is closed under modus ponens, universal generalisation, and first-order substitution:

(MP) if $\vdash_L A$ and $\vdash_L A \supset B$, then $\vdash_L B$,

(UG) if $\vdash_L [t/x]A$, then $\vdash_L \forall xA$,

(Sub) if $\vdash_L A$, then $\vdash_L [y/x]A$.

(See, for example, [Nolt 2021] and [Lambert 2017: 265ff.] for variations.)

Here, $\vdash_L A$ means $A \in L$. $E!x$ abbreviates '$\exists y \, x = y$', where y is the alphabetically first variable other than x. $[t/x]A$ is the result of substituting all occurrences of x in A by t, possibly renaming bound variables to prevent capturing.

Classical logic is obtained by omitting ($\forall$E!) and replacing (FUI) with

(UI) $\forall xA \supset [t/x]A$.

> [Nolt 2021] uses the same axiomatization except that (a) the system is restricted to closed formulas, and (b) (LL) and (Sub) are replaced by the more standard "partial" formulation of Leibniz's Law, which is equivalent to the double-substitution formulation
>
> $$s = t \supset ([s/x]A \supset [t/x]B).$$
>
> I should probably change my axiomatization in accordance with (b). (With my version of (LL), (Sub) is needed to derive $x = y \supset (x = x \supset y = x)$.) I could also adopt (a), at the cost of making comparisons with other treatments in the literature harder.

The minimal normal propositional modal logic K is standardly axiomatized by (Taut), the axiom schema

(K) $\Box(A \supset B) \supset (\Box A \supset \Box B)$

and the "necessitation" rule

(Nec) if $\vdash_L A$, then $\vdash_L \Box A$.

I will use the label 'FK' (for "free K") for the logic that combines the axioms and rules of positive free logic with those of K. Thus FK is the smallest set L that contains all $\mathfrak{L}$-instances of (Taut), (VQ), (UD), (FUI), ($\forall$E!), ($=$R), (LL), and (K) and that is closed under (MP), (UG), and (Nec).

QK (for "quantified K") is the union of classical predicate logic and L, axiomatized by (Taut), (VQ), (UD), (UI), ($=$R), (LL), (K), (MP), (UG), and (Nec).

Standard negative free logic replaces ($=$R) and ($\forall$E!) with ($\forall$$=$R) and (Neg):

($\forall$$=$R) $\quad \forall x(x = x)$,

(Neg) $\quad Px_1 \dots x_n \supset E!x_1 \wedge \dots \wedge E!x_n$.

This system, combined with K, is *not* complete with respect to all single-domain (functional) counterpart structures, where counterpart relations (and assignments) are allowed to be partial. The complete logic of these structures has two further axiom schemas:

(NA) $\quad \neg E!t \supset \Box\neg E!t$,

(TE) $\quad s = t \supset \Box(E!s \supset E!t)$.

I will use 'NK' for the system axiomatized by (Taut), (UD), (VQ), (FUI), (Neg), (LL), ($\forall$$=$R), (K), (NA), (TE), (MP), (UG), and (Nec).

(NA) reflects the fact that non-existent objects don't have any counterparts: if $g(t)$ is undefined and we shift the point of evaluation to another world, then $g(t)$ will still be undefined.

(TE) says that if s is identical to t, and s has a counterpart at some accessible world, then t also has a counterpart at that world. If we had outer domains, an individual could have some existing and some non-existing counterparts at a world, which would render (TE) false.

(NA) should not be confused with the claim that no individual exists at an accessible world that isn't a counterpart of something at the present world. This is rather expressed by the Barcan Formula,

(BF) $\quad \forall x\Box A \supset \Box\forall xA$.

The Barcan Formula is not valid in the class of single-domain models with partial counterpart relations (and partial assignments). For example, if $W = \{w, w'\}$, wRw', $D_w = \emptyset$ and $D_{w'} = \{0\}$, then $\forall x\Box x \neq x$ is true at w but $\Box\forall x \, x \neq x$ is false.

We could provide a semantics for negative modal predicate logics without (NA) and (TE). The relevant models would be dual-domain models in which the extension of all predicates, including identity, is restricted to the inner domain. (NA) then requires that individuals which only figure in the outer domain of a world never have counterparts in the inner domain of another world. (TE) requires that if an individual in the inner domain of a world has a counterpart in the inner domain of another world, then all its counterparts at that world are in its inner domain. The two requirements are obviously independent and non-trivial. Hence the axioms (NA) and (TE) are independent of one another, and of the system that combines negative free logic with K.

3.2 Substitution

The axiomatizations in the previous section involve the substitution operation $[y/x]$. This operation will become important, so let me say a little more about it.

Consider "universal instantiation":

(UI) $\forall x A \supset [t/x]A$.

Informally, we want to allow the inference from $\forall x \exists y Gxy$ to $\exists y Gzy$ but not to $\exists y Gyy$: in the second inference, the variable y that is substituted for x in $\exists y Gxy$ gets "captured" by the quantifier $\exists y$. There are three common strategies to rule out such unwanted instances of (UI).

First, we could require that the substituted term is an individual constant. Constants can't be bound and thus can't be captured. This evidently assumes that the language contains individual constants.

A second strategy is to restrict principles like (UI) to cases where the substituted term t is "free (to be substituted) for x in A", meaning that there is no free occurrence of x in A that falls in the scope of a quantifier binding t.

The third option is to redefine the substitution operation so that it renames bound variables in cases where the substituted term is not free for x in A. This is the option I have chosen, although the others would have worked as well.

I haven't explained how exactly the redefined substitution operation works. There are a number of options. What's important is that the operation satisfies the following

condition, known as the "substitution lemma".

$$g \Vdash [y/x]A \text{ iff } g^{[y/x]} \Vdash A,$$

Here, $g^{[y/x]}$ is the x-variant of g with $g^{[y/x]}(x) = g(y)$. Informally, the substitution lemma ensures that $[y/x]A$ "says about y" what A says about x.

Later, we'll sometimes want to substitute multiple variables at once. I will therefore define substitutions as (total) functions σ from Var to Var. The substitution lemma now requires that

$$g \Vdash \sigma(A) \text{ iff } g \circ \sigma \Vdash A,$$

where $g \circ \sigma$ is the assignment that maps any variable x to $g(\sigma(x))$. The following definition does the job.

Definition 3.1 (Substitution).
A *substitution* (on a set of terms Var) is a total function $\sigma : \text{Var} \rightarrow \text{Var}$. Application of a substitution σ to a formula A is defined as follows.

$$\sigma(Px_1 \ldots x_n) = P\sigma(x_1) \ldots \sigma(x_n)$$

$$\sigma(\neg A) = \neg\sigma(A)$$

$$\sigma(A \supset B) = \sigma(A) \supset \sigma(B)$$

$$\sigma(\forall zA) = \begin{cases} \forall v\sigma^{v \mapsto v}([v/z]A) & \text{if there is an } x \in \text{FV}(\forall zA) \text{ with } \sigma(x) = \sigma(z) \\ \forall\sigma(z)\sigma(A) & \text{otherwise,} \end{cases}$$
where v is the alphabetically first variable not in $\text{FV}(\sigma(A)) \cup \text{FV}(A)$ and $\sigma^{v \mapsto v}$ is the substitution that maps v to v and otherwise coincides with σ.

$$\sigma(\Box A) = \Box\sigma(A).$$

To save space, I will sometimes write 'ϕ^σ' instead of '$\sigma(\phi)$'.

We define $[t/x]$ as the substitution σ that maps x to t and every other variable to itself (Note that $g^{[y/x]} = g \circ [y/x]$.) More generally, $[t_1 \ldots t_n / s_1 \ldots s_n]$ is the substitution σ that maps s_i to t_i (for $1 \le i \le n$) and every other variable to itself.

Lemma 3.1 (Substitution lemma).
For any functional counterpart model $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$, world $w \in W$, assignment g on U_w, $\mathfrak{L}$-formula A, and substitution σ,

$$\mathfrak{M}, w, g \circ \sigma \Vdash \text{ iff } \mathfrak{M}, w, g \Vdash \sigma(A).$$

Proof. By induction on A.

1. A is $Px_1 \ldots x_n$. $\mathfrak{M}, w, g \circ \sigma \Vdash Px_1 \ldots x_n$ iff $(g \circ \sigma)(x_1) \ldots (g \circ \sigma)(x_n) \in I(P)$ (by definition 2.9), iff $g(\sigma(x_1)) \ldots g(\sigma(x_n)) \in I(P)$, iff $\mathfrak{M}, w, g \Vdash P(\sigma(x_1)) \ldots (\sigma(x_n))$ (by definition 2.9), iff $\mathfrak{M}, w, g \Vdash \sigma(Px_1 \ldots x_n)$ (by definition 3.1).

2. A is $\neg B$. $\mathfrak{M}, w, g \circ \sigma \Vdash \neg B$ iff $\mathfrak{M}, w, g \circ \sigma \nVdash B$ by definition 2.9, iff $\mathfrak{M}, w, g \nVdash \sigma(B)$ by induction hypothesis, iff $\mathfrak{M}, w, g \Vdash \neg \sigma(B)$ by definition 2.9, iff $\mathfrak{M}, w, g \Vdash \sigma(\neg B)$ by definition 3.1.

3. A is $B \supset C$. Analogous to the previous case.

4. A is $\forall z B$. By definition 2.9, $\mathfrak{M}, w, g \circ \sigma \Vdash \forall z B$ iff $\mathfrak{M}, w, (g \circ \sigma)^{z \mapsto d} \Vdash B$ for all $d \in D_w$.

   Assume first that there is an $x \in \mathrm{FV}(\forall z B)$ with $\sigma(x) = \sigma(z)$, so that $\sigma(\forall z B) = \forall v \sigma^{v \mapsto v}([v/z]B)$, where $v \notin \mathrm{FV}(B) \cup \mathrm{FV}(\sigma(B))$.

   By definition 2.9, $\mathfrak{M}, w, g \Vdash \forall v \sigma^{v \mapsto v}([v/z]B)$ iff $\mathfrak{M}, w, g^{v \mapsto d} \Vdash \sigma^{v \mapsto v}([v/z]B)$ for all $d \in D_w$. Let d be any element of D_w. By induction hypothesis, $\mathfrak{M}, w, g^{v \mapsto d} \Vdash \sigma^{v \mapsto v}([v/z]B)$ iff $\mathfrak{M}, w, g^{v \mapsto d} \circ \sigma^{v \mapsto v} \Vdash [v/z]B$, iff $\mathfrak{M}, w, g^{v \mapsto d} \circ \sigma^{v \mapsto v} \circ [v/z] \Vdash B$.

   Let g^* be $g^{v \mapsto d} \circ \sigma^{v \mapsto v} \circ [v/z]$. Observe that for any variable x,

   $$g^*(x) = \begin{cases} d & \text{if } (\sigma^{v \mapsto v} \circ [v/z])(x) = v, \\ g(\sigma(x)) & \text{otherwise.} \end{cases}$$

   The first case, $(\sigma^{v \mapsto v} \circ [v/z])(x) = v$, can arise in three ways: (i) $x = z$ or (ii) $x = v$ or (iii) $\sigma(x) = v$. We know that v has no free occurrence in

B. Neither does any variable x for which $\sigma(x) = v$, as otherwise v would be in $FV(\sigma(B))$. (Substitutions never rename free variables. So if $x \in FV(B)$, then $\sigma(x) \in FV(\sigma(B))$.) Cases (ii) or (iii) therefore cannot arise for $x \in FV(B)$. Thus $g^*(x) = (g \circ \sigma)^{z \mapsto d}(x)$ for all $x \in FV(B)$. By the locality lemma 2.2, it follows that $\mathfrak{M}, w, g^* \Vdash B$ iff $\mathfrak{M}, w, (g \circ \sigma)^{z \mapsto d} \Vdash B$.

So we have $\mathfrak{M}, w, g \Vdash \forall v \sigma^{v \mapsto v}([v/z]B)$ iff $\mathfrak{M}, w, (g \circ \sigma)^{z \mapsto d} \Vdash B$ for all $d \in D_w$, iff $\mathfrak{M}, w, g \circ \sigma \Vdash \forall z B$.

Next, assume that there is no $x \in FV(\forall z B)$ with $\sigma(x) = \sigma(z)$, so that $\sigma(\forall z B) = \forall \sigma(z) \sigma(B)$.

By definition 2.9, $\mathfrak{M}, w, g \Vdash \forall \sigma(z) \sigma(B)$ iff $\mathfrak{M}, w, g^{\sigma(z) \mapsto d} \Vdash \sigma(B)$ for all $d \in D_w$. Let d be any element of D_w. By induction hypothesis, $\mathfrak{M}, w, g^{\sigma(z) \mapsto d} \Vdash \sigma(B)$ iff $\mathfrak{M}, w, g^{\sigma(z) \mapsto d} \circ \sigma \Vdash B$. Now for any variable x,

$$(g^{\sigma(z) \mapsto d} \circ \sigma)(x) = \begin{cases} d & \text{if } \sigma(x) = \sigma(z), \\ g(\sigma(x)) & \text{otherwise.} \end{cases}$$

By assumption, there is no variable x besides z in $FV\,B$ for which $\sigma(x) = \sigma(z)$. So $g^{\sigma(z) \mapsto d} \circ \sigma$ and $(g \circ \sigma)^{z \mapsto d}$ agree for all $x \in FV(B)$. As before, it follows that $\mathfrak{M}, w, g \Vdash \forall \sigma(z) \sigma(B)$ iff $\mathfrak{M}, w, g \circ \sigma \Vdash \forall z B$.

5. A is $\Box B$. $\mathfrak{M}, w, g \circ \sigma \Vdash \Box B$ iff $\mathfrak{M}, w', C \circ g \circ \sigma \Vdash B$ for all $C \in K_{w,w'}$ by lemma 2.1, iff $\mathfrak{M}, w', C \circ g \Vdash \sigma(B)$ for all $C \in K_{w,w'}$ by induction hypothesis, iff $\mathfrak{M}, w, g \Vdash \Box \sigma(B)$ by lemma 2.1, iff $\mathfrak{M}, w, g \Vdash \sigma(\Box B)$ by definition 3.1.

The restriction to functional models is crucial. As we'll see in section 7.2, the substitution lemma does not hold in non-functional models.

3.3 Soundness of the base logics

Let's show that all theorems of FK are valid on every functional counterpart structure. As usual, validity means truth at all points of evaluation under all interpretations.

We don't want any (genuinely) empty terms in FK (or QK), so we ignore points of evaluation w, g whose assignment g is not total.

Definition 3.2 (Validity).
An $\mathcal{L}$-formula A is *valid on* a counterpart structure $\mathfrak{S} = \langle W, R, U, D, K \rangle$ iff $\mathfrak{S}, I, w, g \Vdash A$ for all interpretations I on $\mathfrak{S}$, all worlds w in W, and all total assignments g on U_w.

We first prove a lemma about existence.

Lemma 3.2.
For any counterpart model $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$, *world* $w \in W$, *assignment* g *on* U_w, *and term* t,

$$\mathfrak{M}, w, g \Vdash E!t \text{ iff } g(t) \in D_w.$$

Proof. Since $E!t$ is shorthand for $\exists x \, x = t$, definition 2.9 implies that $\mathfrak{M}, w, g \Vdash E!x$ iff there is a $d \in D_w$ for which $\mathfrak{M}, w, g^{x \mapsto d} \Vdash x = t$ and thus $\langle g^{x \mapsto d}(x), g^{x \mapsto d}(t) \rangle \in I_w(=)$. Since x and t are distinct, $g^{x \mapsto d}(x) = d$ and $g^{x \mapsto d}(t) = g(t)$. By definition 2.4, $\langle d, g(x) \rangle \in I_w(=)$ iff $d = g(x)$. So $\mathfrak{M}, w, g \Vdash E!t$ iff there is a $d \in D_w$ for which $d = g(t)$.

Lemma 3.3 (Soundness of the FK axioms).
Every instance of (Taut), (VQ), (UD), (FUI), ($\forall$E!), ($=$R), (LL), *and* (K) *is valid on every total functional counterpart structure.*

Proof. Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ be any total functional counterpart model, w a world in W, and g a (total) assignment on U_w. We show that $\mathfrak{M}, w, g \Vdash A$ for every instance A of every axiom.

1. (Taut). Propositional tautologies are true at every point of evaluation due to the standard rules for $\neg A$ and $A \supset B$ in definition 2.9.

2. (VQ). Suppose for reductio that $\mathfrak{M}, w, g \nVdash A \supset \forall x A$. By definition 2.9, this means that $\mathfrak{M}, w, g \Vdash A$ and $\mathfrak{M}, w, g \nVdash \forall x A$. The latter means that $\mathfrak{M}, w, g^{x \mapsto d} \nVdash A$ for some $d \in D_w$. Since x is not free in A, it follows by the locality lemma 2.2 that $\mathfrak{M}, w, g \nVdash A$. Contradiction.

3. (UD). Assume $\mathfrak{M}, w, g \Vdash \forall x(A \supset B)$ and $\mathfrak{M}, w, g \Vdash \forall x A$. By definition 2.9, then $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A \supset B$ and $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for every $d \in D_w$, wherefore $\mathfrak{M}, w, g^{x \mapsto d} \Vdash B$ for every $d \in D_w$, and so $\mathfrak{M}, w, g \Vdash \forall x B$.

4. (FUI). Assume $\mathfrak{M}, w, g \Vdash \forall x A$ and $\mathfrak{M}, w, g \Vdash E!t$. By lemma 3.2, the latter means that $g(t) \in D_w$. By definition 2.9, the former means that $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for every $d \in D_w$. So in particular, $\mathfrak{M}, w, g^{x \mapsto g(t)} \Vdash A$. Since $g^{x \mapsto g(t)} = g^{[t/x]}$, it follows by lemma 3.1 that $\mathfrak{M}, w, g \Vdash [t/x]A$.

5. ($\forall$E!). $\mathfrak{M}, w, g \Vdash \forall x E!x$ iff $\mathfrak{M}, w, g^{x \mapsto d} \Vdash E!x$ for every $d \in D_w$ by definition 2.9, iff $g^{x \mapsto d}(x) \in D_w$ for every $d \in D_w$ by lemma 3.2. Since $g^{x \mapsto d}(x) = d$, this is always the case.

6. (=R). By definition 2.4, $I_w = \{\langle d, d \rangle : d \in U_w\}$. Since $g(t) \in U_w$, it follows by definition 2.9 that $\mathfrak{M}, w, g \Vdash t = t$.

7. (LL). Assume $\mathfrak{M}, w, g \Vdash s = t$ and $\mathfrak{M}, w, g \Vdash A$. By definitions 2.9 and 2.4, the former implies that $g(s) = g(t)$. So $g^{[s/t]} = g$. Since $\mathfrak{M}, w, g \Vdash A$, we have $\mathfrak{M}, w, g^{[s/t]} \Vdash A$. By the substitution lemma 3.1, it follows that $\mathfrak{M}, w, g \Vdash [s/t]A$.

8. (K). Assume $\mathfrak{M}, w, g \Vdash \Box(A \supset B)$ and $\mathfrak{M}, w, g \Vdash \Box A$. By definition 2.9, this means that for all w', g' such that $w, g \rhd w', g'$, $\mathfrak{M}, w', g' \Vdash A \supset B$ and $\mathfrak{M}, w', g' \Vdash A$ It follows by the clause for $\supset$ in definition 2.9 that $\mathfrak{M}, w', g' \Vdash B$ for all such w', g'. And so $\mathfrak{M}, w, g \Vdash \Box B$ by the clause for the box.

Lemma 3.4 (Soundness of the FK rules).
If all elements of a set of formulas Γ are valid on a counterpart structure $\mathfrak{S}$, and Γ is extended by (MP)*,* (UG)*,* (Nec)*, or* (Sub)*, then the new sentences are still valid on $\mathfrak{S}$.*

Proof.
1. (MP). Assume that A and $A \supset B$ are in Γ. By definition 2.9, this means that $\mathfrak{M}, w, g \Vdash A$ and $\mathfrak{M}, w, g \Vdash A \supset B$ for every $\mathfrak{M}, w, g$ based on $\mathfrak{S}$. Then $\mathfrak{M}, w, g \Vdash B$ for every such $\mathfrak{M}, w, g$, by definition 2.9.

2. (UG). We argue by contraposition. Assume that $\mathfrak{M}, w, g \nVdash \forall x A$ for some $\mathfrak{M}, w, g$ based on $\mathfrak{S}$. Then $\mathfrak{M}, w, g^{x \mapsto d} \nVdash A$ for some $d \in D_w$, by definition 2.9. So A is not valid on $\mathfrak{S}$.

3. (Nec). Again, we argue by contraposition. Assume that $\mathfrak{M}, w, g \nVdash \Box A$ for some $\mathfrak{M}, w, g$ based on $\mathfrak{S}$. Then $\mathfrak{M}, w', g' \nVdash A$ for some w', g' with $w, g \vartriangleright w', g'$, by definition 2.9. So A is not valid on $\mathfrak{S}$.

4. (Sub). For contraposition, assume that $\mathfrak{M}, w, g \nVdash \sigma(A)$ for some $\mathfrak{M}, w, g$ based on $\mathfrak{S}$. By the substitution lemma 3.1, $\mathfrak{M}, w, g \circ \sigma \nVdash A$. Since $g \circ \sigma$ is an assignment on U_w, this means that A is not valid in Σ.

Theorem 3.5 (Soundness of FK).
Every member of FK is valid on every total functional counterpart structure.

Proof. Immediate from lemmas 3.3 and 3.4.

Lemma 3.6 (Soundness of (UI)).
Every instance of (UI) *is valid on every classical counterpart structure.*

Proof. Recall that a structure is classical if it is total, functional, and single-domain, so that $D_w = U_w$ for all $w \in W$.

Assume $\mathfrak{M}, w, g \Vdash \forall x A$. By definition 2.9, $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for every $d \in D_w$. Since g is total, $g(t) \in U_w$. So $g(t) \in D_w$. And so $\mathfrak{M}, w, g^{x \mapsto g(t)} \Vdash A$. Since $g^{x \mapsto g(t)} = g^{[t/x]}$, it follows by lemma 3.1 that $\mathfrak{M}, w, g \Vdash [t/x]A$.

Theorem 3.7 (Soundness of QK).
Every member of QK *is valid on every classical counterpart structure.*

Proof. Immediate from lemmas 3.3, 3.4 and 3.6.

For negative logics, we redefine the concept of validity. We want to allow for terms that go genuinely empty. So assignment functions can be partial.

Definition 3.3 (N-Validity).
An $\mathfrak{L}$-formula A is *n-valid on* a counterpart structure $\mathfrak{S} = \langle W, R, U, D, K \rangle$ iff $\mathfrak{S}, I, w, g \Vdash A$ for all interpretations I on $\mathfrak{S}$, all worlds w in W, and all partial assignments g on U_w.

Lemma 3.8 (Soundness of the NK axioms).
Every instance of (Taut), (VQ), (UD), (FUI), (Neg), ($\forall = R$), (LL), (NA), (TE), *and* (K) *is N-valid on every single-domain functional counterpart structure.*

Proof. Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ be any single-domain functional counterpart model, w a world in W, and g an assignment on U_w.

We show that $\mathfrak{M}, w, g \Vdash A$ for every instance A of every axiom. The proofs for (Taut), (VQ), (UD), (FUI), (LL), and (K) are just as in the proof of lemma 3.3. Let's go through the remaining cases.

- (Neg). Assume $\mathfrak{M}, w, g \Vdash Px_1 \ldots x_n$. By definition 2.4, $I(P) \subseteq U_w^n$. Since $D = U$, we have $I(P) \subseteq D_w^n$. By definition 2.9, $\langle g(x_1), \ldots, g(x_n) \rangle \in I(P)$. So

$g(x_i) \in D_w$ for all $x_i \in x_1, \ldots, x_n$. So $\mathfrak{M}, w, g \Vdash E!x_1 \wedge \ldots \wedge E!x_n$ by lemma 3.2.

- $(\forall = \mathrm{R})$. By definition 2.9, $\mathfrak{M}, w, g \Vdash \forall x\ x = x$ iff $\mathfrak{M}, w, g^{x \mapsto d} \Vdash x = x$ for all $d \in D_w$, which in turns holds iff $\langle g^{x \mapsto d}(x), g^{x \mapsto d}(x) \rangle \in I(=)$ for all $d \in D_w$. This is always the case by definition 2.4.

- (NA). Assume $\mathfrak{M}, w, g \Vdash \neg E!x$. By definition 2.9, this means that $g(x) \notin D_w$. Since $D = U$, $g(x)$ must be undefined. In that case, there is no world w', individual d and counterpart relation $C \in K_{w,w'}$ such that $g(x)Cd$. Thus whenever $w, g \rhd w', g'$, then $g'(x)$ is undefined, by definition 2.8, By lemma 3.2, $\mathfrak{M}, w', g' \Vdash \Box \neg E!x$ if $g'(x)$ is undefined. So $\mathfrak{M}, w, g \Vdash \Box \neg E!x$ by definition 2.9.

- (TE). Assume $\mathfrak{M}, w, g \Vdash x = y$. Then $g(x) = g(y)$ by definitions 2.4 and 2.9. Let w', g' be such that $w, g \rhd w', g'$ and $\mathfrak{M}, w', g' \Vdash E!x$. We have $g'(x) \in D_{w'}$ by lemma 3.2. By definition 2.8, $w, g \rhd w', g'$ means that there is a $C \in K_{w,w'}$ such that $g(x)Cg'(x)$. So there is a $d \in D_{w'}$ (namely, $g'(x)$) for which $g(y)Cd$. By definition 2.8, it follows that $g'(y)$ can't be undefined. So $g'(y) \in U_{w'}$, and so $g'(y) \in D_{w'}$ because $U = D$. By lemma 3.2, it follows that $w', g' \Vdash E!y$.

Theorem 3.9 (Soundness of NK**).**
Every member of NK *is N-valid on every single-domain functional counterpart structure.*

Proof. Immediate from lemmas 3.8 and 3.4.

3.4 Some consequences

In this section, I will prove a few consequences of the above axiomatizations. Some of these will be needed in the completeness proofs. The consequences hold not only

for the base logics, but also for extensions of the base logics. We define two types of extension.

Definition 3.4 (Positive logics).
A set of $\mathfrak{L}$-sentences is a *positive logic* if it includes FK and is closed under (MP), (UG), (Nec) and (Sub).

Definition 3.5 (Negative logics).
A set of $\mathfrak{L}$-sentences is a *negative logic* if it includes NK and is closed under (MP), (UG), (Nec) and (Sub).

These definitions allow for "logics" in which (say) F_1x is a theorem but not F_2x. A genuine logic should also satisfy some second-order closure condition, but we will not worry about that here.

From now on, let L be an arbitrary positive or negative logic, in the sense of the above definitions.

Lemma 3.10 (Closure under propositional consequence).
For all $\mathfrak{L}$-formulas $A_1, \ldots, A_n, B$,

(PC) *if $\vdash_L A_1$, ..., $\vdash_L A_n$, and B is a propositional consequence of $A_1, \ldots, A_n$, then $\vdash_L B$.*

Proof. If B is a propositional consequence of $A_1, \ldots, A_n$, then $A_1 \supset (\ldots \supset (A_n \supset B) \ldots)$ is a tautology. So by (Taut), $\vdash_L A_1 \supset (\ldots \supset (A_n \supset B) \ldots)$. If $\vdash_L A_1, \ldots, \vdash_L A_n$, then by n applications of (MP), $\vdash_L B$.

When giving proofs, I will often omit reference to (PC).

Lemma 3.11 (Redundant axioms).
For any $\mathfrak{L}$-formulas A and variables x,

($\forall$E!) $\quad \vdash_L \forall x E!x,$

($\forall = $R) $\quad \vdash_L \forall x (x = x).$

Proof. If L is positive, then ($\forall$E!) is an axiom. In N, we have $\vdash_L x = x \supset E!x$ by (Neg); so by (UG) and (UD), $\vdash_L \forall x(x = x) \supset \forall x E!x$. Since $\vdash_L \forall x(x = x)$ by (=R), $\vdash_L \forall x E!x$.

If L is negative, then ($\forall$ = R) is an axiom. In P, we have $\vdash_L x = x$ by (=R), and so ($\forall$ = R) by (UG).

Lemma 3.12 (Existence and self-identity).
If L is negative*, then for any $\mathfrak{L}$-variable x,*

(EI) $\vdash_L E!x \leftrightarrow x = x;$

Proof. By (FUI), $\vdash_L \forall x(x = x) \supset (E!x \supset x = x)$. By ($\forall$ = R), $\vdash_L \forall x(x = x)$. So $\vdash_L E!x \supset x = x$. Conversely, $x = x \supset E!x$ by (Neg).

Lemma 3.13 (Symmetry and transitivity of identity).
For any $\mathfrak{L}$-variables x, y, z,

(= S) $\vdash_L x = y \supset y = x;$

(= T) $\vdash_L x = y \supset y = z \supset x = z.$

Proof. For (= S), let v be some variable $\notin \{x, y\}$. Then

1. $\vdash_L v = y \supset (v = x \supset y = x)$. (LL)
2. $\vdash_L x = y \supset (x = x \supset y = x)$. (1, (Sub))
3. $\vdash_L x = y \supset x = x$. ((=R), or (Neg) and ($\forall$ = R))
4. $\vdash_L x = y \supset y = x$. (2, 3)

For (= T),

$$
\begin{array}{lll}
1. & \vdash_L x = y \supset y = x. & (=S) \\
2. & \vdash_L y = x \supset (y = z \supset x = z). & (LL) \\
3. & \vdash_L x = y \supset (y = z \supset x = z). & (1, 2)
\end{array}
$$

Lemma 3.14 (Necessity of identity).
For any $\mathcal{L}$-variables x, y,

(NI) $x = y \supset \Box (x = x \supset x = y).$

Proof. Let v be some variable $\notin \{x, y\}$. Then

$$
\begin{array}{lll}
1. & \vdash_L v = y \supset \Box (x = x \supset x = v) \supset \Box (x = x \supset x = y). & (LL) \\
2. & \vdash_L x = y \supset \Box (x = x \supset x = x) \supset \Box (x = x \supset x = y). & (1, (Sub)) \\
3. & \vdash_L x = x \supset x = x. & (Taut) \\
4. & \vdash_L \Box (x = x \supset x = x). & (3, (Nec)) \\
5. & \vdash_L x = y \supset \Box (x = x \supset x = y). & (2, 4)
\end{array}
$$

Finally, we prove that sentences that differ by renaming bound variables are provably equivalent.

Definition 3.6 (Alphabetic variant).
A formula A' is an *alphabetic variant of* a formula A if one of the following conditions is satisfied.

1. $A = A'$.

2. $A = \neg B$, $A' = \neg B'$, and B' is an alphabetic variant of B.

3. $A = B \supset C$, $A' = B' \supset C'$, and B', C' are alphabetic variants of B, C, respectively.

4. $A = \forall x B$, $A' = \forall z[z/x]B'$, B' is an alphabetic variant of B, and either $z = x$ or $z \notin \mathrm{Var}(B')$.

5. $A = \Box B$, $A' = \Box B'$, and B' is an alphabetic variant of A'.

Lemma 3.15 (Syntactic alpha-conversion).
If A, A' are $\mathfrak{L}$-formulas, and A' is an alphabetic variant of A, then

(AC) $\vdash_L A \leftrightarrow A'$.

Proof. by induction on A.

1. A is atomic. Then $A = A'$ and $A \leftrightarrow A'$ is a propositional tautology.

2. A is $\neg B$. Then A' is $\neg B'$, where B' is an alphabetic variant of A'. By induction hypothesis, $\vdash_L B \leftrightarrow B'$. So by (PC), $\vdash_L \neg B \leftrightarrow \neg B'$.

3. A is $B \supset C$. Then A' is $B' \supset C'$, where B', C' are alphabetic variants of B, C, respectively. By induction hypothesis, $\vdash_L B \leftrightarrow B'$ and $\vdash_L C \leftrightarrow C'$. So by (PC), $\vdash_L (B \supset C) \leftrightarrow (B' \supset C')$.

4. A is $\forall x B$. Then A' is either $\forall x B'$ or $\forall z[z/x]B'$, where B' is an alphabetic variant of B and $z \notin \mathrm{Var}(B')$. Assume first that A' is $\forall x B'$. By induction hypothesis, $\vdash_L B \leftrightarrow B'$. So by (UG) and (UD), $\vdash_L \forall x B \leftrightarrow \forall x B'$.

   Alternatively, assume B is $\forall z[z/x]B'$ and $z \notin \mathrm{Var}(B')$. Since B' differs from B at most in renaming bound variables, if z were free in B, then

$z \in \text{Var}(B')$. So z is not free in B. Then

1.	$\vdash_L B \leftrightarrow B'$	(induction hypothesis)
2.	$\vdash_L [z/x]B \leftrightarrow [z/x]B'$	(1, (Sub))
3.	$\vdash_L \forall x B \supset Ez \supset [z/x]B$	(FUI)
4.	$\vdash_L \forall x B \supset Ez \supset [z/x]B'$	(2, 3)
5.	$\vdash_L \forall z \forall x B \supset \forall z Ez \supset \forall z[z/x]B'$	(4, (UG), (UD))
6.	$\vdash_L \forall z Ez$	($\forall$Ex)
7.	$\vdash_L \forall z \forall x B \supset \forall z[z/x]B'$	(5, 6)
8.	$\vdash_L \forall x B \supset \forall z \forall x B$	((VQ), z not free in B)
9.	$\vdash_L \forall x B \supset \forall z[z/x]B'.$	(7, 8)

Conversely,

10.	$\vdash_L \forall z[z/x]B' \supset Ex \supset [x/z][z/x]B'$	(FUI)
11.	$\vdash_L \forall z[z/x]B' \supset Ex \supset B$	(1, 10, $z \notin \text{Var}(B')$)
12.	$\vdash_L \forall x \forall z[z/x]B' \supset \forall x B$	(11, (UG), (UD), ($\forall$E!x))
13.	$\vdash_L \forall z[z/x]B' \supset \forall x \forall z[z/x]B'$	(VQ)
14.	$\vdash_L \forall z[z/x]B' \supset \forall x B$	(12, 13)

5. A is $\Box B$. Then A' is $\Box B'$, where B' is an alphabetic variant of B. By induction hypothesis, $\vdash_L B \leftrightarrow B'$. Then by (Nec), $\vdash_L \Box(B \leftrightarrow B')$, and by (K) and (PC), $\vdash_L \Box B \leftrightarrow \Box B'$.

4 Canonical models

4.1 Preview

We are going to use the canonical model technique to prove that our base logics are complete. Recall that a logic L is *(strongly) complete* with respect to a class of structures Σ iff every L-consistent set of formulas Γ is satisfiable in Σ. A set of formulas Γ is *L-consistent* iff there are no formulas $A_1, \ldots, A_n$ in Γ for which $\vdash_L \neg(A_1 \wedge \cdots \wedge A_n)$. In our context, a set of formulas Γ is *satisfiable in* a class of structures iff there is a structure $\mathfrak{S} \in \Sigma$, an interpretation I on $\mathfrak{S}$, a world w in $\mathfrak{S}$, and a (suitable) assignment g on U_w such that $\mathfrak{S}, I, w, g \Vdash A$ for all formulas A in Γ.

To establish completeness for a logic L, we define a *canonical model* $\mathfrak{M}_L$ whose worlds are maximal L-consistent sets of formulas. We show that for each world w there is an assignment g_w such that a formula is true at w under g_w iff it is a member of w:

$$\mathfrak{M}_L, w, g_w \models A \text{ iff } A \in w.$$

This is known as the *truth lemma*. Since every L-consistent set of formulas can be extended to a maximal L-consistent set, it follows that every L-consistent set of formulas is satisfiable in any set of structures that contains the structure of the canonical model.

To secure the truth lemma for atomic sentences, we stipulate that for any variable x, $g_w(x)$ is the set of variables z for which w contains $x = z$. (We denote this set by '$[x]_w$'.) The $\mathfrak{M}_L$-interpretation I will then assign to each predicate P at w the set of n-tuples $\langle [x_1]_w, \ldots, [x_n] \rangle$ for which $Px_1 \ldots x_n \in w$.

When we turn to quantified sentences, we face a well-known problem. Classical first-order logic does not require every individual to have a name: there are consistent sets Γ that contain $\exists x F x$ as well as $\neg F v_i$ for every variable v_i. If w extends such a set, then the construction just outlined would make $I_w(F)$ the empty set. We would have $\mathfrak{M}, w, g_w \Vdash \neg \exists x F x$ even though $\exists x F x \in w$. The standard response, which

we adopt, is to stipulate that the worlds in a canonical model are *witnessed*, so that whenever an existential formula $\exists x F x$ is in w, then some witnessing instance $F y$ is in w as well. To ensure that sets like Γ are satisfiable in $\mathfrak{M}_L$, we construct the worlds of $\mathfrak{M}_L$ in a larger language $\mathfrak{L}^*$ that adds infinitely many new variables to the original language $\mathfrak{L}$.

In modal predicate logic, the problem of witnesses reappears in another form. If the worlds of a canonical model are maximal consistent and witnessed sets of $\mathfrak{L}^*$-sentences, then a world w might contain $\Diamond \exists x F x$ as well as $\Box \neg F v_i$ for every $\mathfrak{L}^*$-variable v_i. For these sentences to be true at w under g_w, Kripke semantics demands that there is a world w' accessible from w that verifies all instances of $\neg F v_i$ as well as $\exists x F x$. But then w' isn't witnessed!

In response, one can use different extensions of $\mathfrak{L}_w$ for different worlds w, and stipulate that if wRw' then $\mathfrak{L}_w \subset \mathfrak{L}_{w'}$. (See [Hughes and Cresswell 1996: ch.15], [Corsi 2002b].) This, however, does not work for quantified extensions of B, where the accessibility relation is symmetrical, as L_w can't be a proper subset of itself. For extensions of B with a classical theory of quantification, one can, however, construct all worlds in the same language $\mathfrak{L}^*$. That's because these logics contain the Barcan Formula, which ensures that any world that contains $\Diamond \exists x F x$ also contains $\exists x \Diamond F x$, so the standard witnessing requirement ensures that it contains $\Diamond F y$ for some y. (See [Hughes and Cresswell 1996: ch.14].) A similar approach works if the underlying predicate logic is free, provided that it is closed under what Corsi [2002b] calls the "Extended Barcan Rule" (see also [Hughes and Cresswell 1996: ch.16]).

Counterpart semantics provides an easier way out. By definition 2.9, the truth of $\Diamond \exists x F x$ at w under g_w requires that there is an accessible world w' for which $I_{w'}(F)$ is non-empty. But $\Box \neg F v_i$ at w under g_w only requires that $\neg F v_i$ is true at w' under all assignments g' that map each variable to a *counterpart* of its g_w-value. If, say, each $[x_i]_w$ at w has $[x_{2i}]_{w'}$ as its counterpart at w', then $[x_1]_{w'}$, $[x_3]_{w'}$, etc. might become available as witnesses for $\exists x F x$.

Fixing a particular counterpart relation like this would make it hard to prove completeness for stronger systems. A better idea – due to [Kutz 2000] – is to determine which individuals are counterparts of one another by looking at what the relevant worlds say about them.

For example, suppose w and w' look as follows.

$$w : \{x \neq y, \Box x \neq y, \Box Fx, \Box Fy, \dots\}$$
$$w' : \{\neg Fx, Fu, Fv, u \neq v, \dots\}$$

We can see that $[x]_{w'}$ does not qualify as counterpart of $[x]_w$, since it doesn't satisfy the "modal profile" that w attributes to $[x]_w$: all counterparts of $[x]_w$ should be F. Both $[u]_{w'}$ and $[v]_{w'}$ meet this condition. We might say that both of them are counterparts of $[x]_w$. But then they should also be counterparts of $[y]_w$, and we get a violation of the "joint modal profile" expressed by $\Box x \neq y$, which requires that no counterpart of $[x]_w$ is identical to any counterpart of $[y]_w$.

Structurally, this is Hazen's problem of internal relations (discussed in section 2.2). In response, we assume that there can be multiple counterpart relations. One (C_1) links $[x]_w$ uniquely with $[u]_{w'}$, and $[y]_w$ with $[v]_w$, another (C_2) links $[x]_w$ with $[v]_{w'}$, and $[y]_w$ with $[u]_{w'}$. So $[x]_w$ has both $[u]_{w'}$ and $[v]_{w'}$ as counterpart, but the pair $\langle [x]_w, [y]_w \rangle$ has only two rather than four counterparts: $\langle [u]_{w'}, [v]_{w'} \rangle$ relative to C_1 and $\langle [v]_{w'}, [u]_{w'} \rangle$ relative to C_2.

In general, we'll say that a world w' in the canonical model is accessible from a world w iff there is a substitution σ such that w' contains all sentences $\sigma(A)$ for which w contains $\Box A$. We'll say that C is a counterpart relation in $K_{w,w'}$ iff it links individuals that are related by some such substitution, so that $[x]_w C [x']_{w'}$ iff there is a $y \in [x]_w$ with $\sigma(y) \in [x']_{w'}$.

On to the details.

4.2 Constructing canonical models

Let L be a positive or negative modal predicate logic (in the sense of section 3.4). Let $\mathfrak{L}^*$ be generated by the rules of definition 2.3 applied to the variables $\mathrm{Var}^* = \mathrm{Var} \cup \mathrm{Var}^+$, where Var^+ is a denumerable set of new variables. The worlds in the canonical model for L are the Henkin sets for L in $\mathfrak{L}^*$, defined as follows.

Definition 4.1 (Henkin set).
A *Henkin set* for L in $\mathfrak{L}^*$ is a set H of $\mathfrak{L}^*$-formulas that is

1. *L-consistent*: there are no $A_1, \dots, A_n \in H$ with $\vdash_L \neg(A_1 \wedge \dots \wedge A_n)$,

2. *maximal*: for every $\mathfrak{L}^*$-formula A, H contains either A or $\neg A$, and

3. *witnessed*: whenever H contains an existential formula $\exists xA$, then there is a variable $y \notin \mathrm{FV}(A)$ such that H contains $[y/x]A$ as well as $E!y$.

I write $\mathfrak{H}_L$ for the class of Henkin sets for L in $\mathfrak{L}^*$.

Definition 4.2 (Variable classes).
For any Henkin set H, define $\sim_H$ to be the relation on Var^* such that $x \sim_H y$ iff $x = y \in H$. For any variable x, let $[x]_H$ be $\{y : x \sim_H y\}$.

This definition is justified by the following lemma.

Lemma 4.1 ($\sim$-Lemma).
$\sim_H$ *is an equivalence relation on the set* $\{x : x = x \in H\}$.

Proof. Immediate from lemma 3.13.

Definition 4.3 (Accessibility via transformations).
Let w, w' be Henkin sets and σ a substitution. We say that w' *is accessible from w via τ* (for short: $w \xrightarrow{\tau} w'$) iff $\{\sigma(X) : \Box X \in w\} \subseteq w'$.

Definition 4.4 (Canonical model).
The *canonical model* $\langle W, R, U, D, K, I \rangle$ for a logic L is defined as follows.

1. The *worlds W* are the Henkin sets $\mathfrak{H}_L$.

2. For each $w \in W$, the *outer domain U_w* comprises all non-empty sets $[x]_w$ for x in Var^*.

3. For each $w \in W$, the *inner domain D_w* comprises all sets $[x]_w$ for which $E!x \in w$. That is, $D_w = \{[x]_w : E!x \in w\}$.

4. The *accessibility relation R* holds between world w and world w' iff there is a substitution σ such that $w \xrightarrow{\sigma} w'$.

5. C is a *counterpart relation* in $K_{w,w'}$ iff there is a substitution σ such that (i) $w \xrightarrow{\sigma} w'$ and (ii) for all $d \in U_w, d' \in U_{w'}, dCd'$ iff there is an $x \in d$ with $\sigma(x) \in d'$.

6. The *interpretation I* assigns to every non-logical predicate P and world w the set $I_w(P) = \{\langle [x_1]_w, \dots, [x_n]_w \rangle : Px_1 \dots x_n \in w\}$.

The term '$\{\langle [x_1]_w, \dots, [x_n]_w \rangle : Px_1 \dots x_n \in w\}$' in clause 6 is meant to denote the set of n-tuples $\langle d_1, \dots, d_n \rangle$ for which there are variables $x_1, \dots, x_n$ such that $d_i = [x_i]_w$ (for $1 \le i \le n$) and $Px_1 \dots x_n \in w$. These d_i are guaranteed to be non-empty because $x_i = x_i \in w$ whenever $Px_1 \dots x_n \in w$: if L is positive, then $\vdash_L z_i = z_i$ by (=R); if L is negative, then $\vdash_L Pz_1 \dots z_n \supset Ez_i$ by (Neg) and hence $\vdash_L Pz_1 \dots z_n \supset z_i = z_i$ by ($\forall$=R) and (FUI).

Definition 4.5 (Canonical Assignment).

If w is a world in a canonical model $\mathfrak{M}$ then the *canonical assignment for w* is the function g_w that maps every $x \in \mathrm{Var}^*$ for which $[x]_w$ is non-empty to $[x]_w$.

If L is a positive logic then $[x]_w$ is never empty, since $\vdash_L x = x$.

Lemma 4.2 (Functionality of canonical models).
The canonical model of every positive or negative logic is functional.

Proof. We have to show (by definition 2.2) that any $C \in K_{w,w'}$ relates each $d \in U_w$ to at most one $d' \in U_{w'}$. So let w, w' be any Henkin sets and $C \in K_{w,w'}$. By definition 4.4, this means that there is a substitution σ such that (i) $w \xrightarrow{\sigma} w'$ and (ii) for all $d \in U_w, d' \in U_{w'}, dCd'$ iff there is a $x \in d'$ with $\sigma(x) \in d'$. Assume dCd' and dCd''. So there are $x, y \in d$ with $\sigma(x) \in d'$ and $\sigma(y) \in d''$. From $x, y \in d$ we have $x = y \in w$, by definitions 4.4 and 4.2. From $\sigma(x) \in d'$ we likewise have $\sigma(x) = \sigma(x) \in w'$. By (NI), w contains

$x = y \supset \Box(x = x \supset x = y)$. So w contains $\Box(x = x \supset x = y)$. Since $w \xrightarrow{\sigma} w'$, w' contains $\sigma(x) = \sigma(x) \supset \sigma(x) = \sigma(y)$. So w' contains $\sigma(x) = \sigma(y)$. And so $d' = d''$.

Lemma 4.3 (Charge of canonical models).

If L is positive then the canonical model for L is total. If L is negative then the canonical model for L is single-domain.

Proof. If L is positive then for all $\mathcal{L}^*$-variables x, every Henkin set for L contains $x = x$ (by (=R)). So $[x]_w$ is never empty, and nor is $[x^\tau]_{w'}$, for any world w' with $w \xrightarrow{\sigma} w'$. So any $C \in K_{w,w'}$ relates every member of U_w to at least one member of $U_{w'}$.

If L is negative then every Henkin set for L contains $x = x \supset Ex$, for all $\mathcal{L}^*$-variables x (by (Neg)). So $[x]_w \neq \emptyset$ iff $E!x \in w$. So $D_w = U_w$ for all worlds w.

Lemma 4.4 (Extensibility Lemma).

If Γ is an L-consistent set of $\mathcal{L}^*$-sentences in which infinitely many $\mathcal{L}^*$-variables do not occur, then there is a Henkin set $H \in \mathfrak{H}_L$ such that $\Gamma \subseteq H$.

Proof. Let $S_1, S_2, \ldots$ be an enumeration of all $\mathcal{L}^*$-sentences, and $z_1, z_2, \ldots$ an enumeration of the unused $\mathcal{L}^*$-variables in such a way that $z_i \notin \mathrm{FV}(S_1 \wedge \ldots \wedge S_i)$. Let $\Gamma_0 = \Gamma$, and define Γ_n for $n \geq 1$ as follows.

   (i) If $\Gamma_{n-1} \cup \{S_n\}$ is not L-consistent, then $\Gamma_n = \Gamma_{n-1}$;

  (ii) else if S_n is an existential formula $\exists x A$, then $\Gamma_n = \Gamma_{n-1} \cup \{\exists x A, [z_n/x]A, Ez_n\}$;

 (iii) else $\Gamma_n = \Gamma_{n-1} \cup \{S_n\}$.

Define H as the union of all Γ_n. We show that H is a Henkin set for L.

First, H is L-consistent. We know that Γ_0 is L-consistent. We show that whenever Γ_{n-1} is L-consistent then so is Γ_n. It follows that every finite subset

of H is L-consistent and hence that H itself is L-consistent. So assume (for $n > 0$) that Γ_{n-1} is L-consistent. Then Γ_n is constructed by applying one of (i)–(iii).

If case (i) in the construction applies, then $\Gamma_n = \Gamma_{n-1}$, and so Γ_n is also L-consistent.

Assume case (ii) in the construction applies. So $\Gamma_n = \Gamma_{n-1} \cup \{\exists x A, [z_n/x]A, Ez\}$. If Γ_n is not L-consistent then there is a finite subset $\{C_1, \ldots, C_m\} \subseteq \Gamma_{n-1}$ such that

1. $\vdash_L \neg(C_1 \wedge \ldots \wedge C_m \wedge \exists x A \wedge [z_n/x]A \wedge Ez_n)$.

Let $\underline{C}$ abbreviate $C_1 \wedge \ldots \wedge C_m$. Then

2. $\vdash_L \underline{C} \wedge \exists x A \supset (Ez_n \supset \neg[z_n/x]A)$.   (1)
3. $\vdash_L \forall z_n(\underline{C} \wedge \exists x A) \supset \forall z_n Ez_n \supset \forall z_n \neg[z_n/x]A$.   (2, (UG), (UD))
4. $\vdash_L \underline{C} \wedge \exists x A \supset \forall z_n(\underline{C} \wedge \exists x A)$.   ((VQ), z_n not free in Γ_{n-1})
5. $\vdash_L \underline{C} \wedge \exists x A \supset \forall z_n Ez_n \supset \forall z_n \neg[z_n/x]A$.   (3, 4)
6. $\vdash_L \underline{C} \wedge \exists x A \supset \forall z_n \neg[z_n/x]A$.   (5, ($\forall$E!))
7. $\vdash_L \forall z_n \neg[z_n/x]A \leftrightarrow \forall x \neg A$.   ((AC), $z_n \notin \mathrm{Var}(A)$)
8. $\vdash_L \underline{C} \wedge \exists x A \supset \neg \exists x A$.   (6, 7)

So $\{C_1, \ldots C_m, \exists x A\}$ is not L-consistent, contradicting the assumption that clause (ii) applies.

If case (iii) in the construction applies then $\Gamma_n = \Gamma_{n-1} \cup \{S_n\}$ is L-consistent, as otherwise case (i) would have applied.

Next, we have to show that H is maximal. Assume neither S_n nor $\neg S_n$ is in H. Then case (i) applied for both, meaning that neither $\Gamma_{n-1} \cup \{S_n\}$ nor $\Gamma_{n-1} \cup \{\neg S_n\}$ is L-consistent. So there are $C_1, \ldots, C_m, D_1, \ldots, D_k \in \Gamma_{n-1}$ such that $\vdash_L C_1 \wedge \ldots \wedge C_m \supset \neg S_n$ and $\vdash_L D_1 \wedge \ldots \wedge D_k \supset \neg\neg S_n$. By (PC), it follows that

$$\vdash_L C_1 \wedge \ldots \wedge C_m \wedge D_1 \wedge \ldots \wedge D_k \supset (\neg S_n \wedge \neg\neg S_n).$$

But then H is inconsistent, contradicting what was just shown.

Finally, we have to show that H is witnessed. This is guaranteed by clause (ii) of the construction and the fact that $z_n \notin \mathrm{FV}(S_n)$.

Now we want to show that a sentence is true at a world in a canonical model relative to the canonical assignment iff it is an element of the world.

A minor complication arises from the fact that quantifiers and modal operators can shift the assignment function away from the canonical assignment. For example, we have $\mathfrak{M}, w, g_w \models \Box A$ iff $\mathfrak{M}, w', g' \models A$ for all w', g' with $w, g_w \rhd w', g'$, and g' may not be the canonical assignment for w'. So we can't assume, by induction hypothesis, that A is an element of w'. The following lemma helps get around this problem.

Lemma 4.5.

If $\mathfrak{M}$ is the canonical model for some positive or negative logic, w is a world in $\mathfrak{M}$, and A is any sentence, then

$$\mathfrak{M}, w, g_w \Vdash \Box A \quad \textit{iff} \quad \mathfrak{M}, w', g_{w'} \circ \sigma \Vdash A \textit{ for all } w', \sigma \textit{ with } w \xrightarrow{\sigma} w'.$$

Proof. We know from lemma 4.2 that $\mathfrak{M}$ is functional. By lemma 2.1, $\mathfrak{M}, w, g_w \Vdash \Box A$ iff $\mathfrak{M}, w', C \circ g_w \Vdash A$ for all w', C with wRw' and $C \in K_{w,w'}$. By definition 4.4, wRw' iff there is a substitution σ with $w \xrightarrow{\sigma} w'$, and $C \in K_{w,w'}$ iff there is some such substitution such that for all $v, v' \in \mathrm{Var}^*$, $[v]_w C[v']_{w'}$ iff there is an $x \in [v]_w$ with $\sigma(x) \in [v']_{w'}$.

Note that if $x \in [v]_w$ and $\sigma(x) \in [v']_{w'}$ then $\sigma(x') \in [v']_{w'}$ for every $x' \in [v]_w$. For suppose $x' \in [v]_w$. Then $x = x' \in w$, and so $\Box x = x' \in w$ by (NI), and so $\sigma(x) = \sigma(x') \in w'$ by the fact that $w \xrightarrow{\sigma} w'$. Since $\sigma(x) \in [v']_{w'}$, we therefore have $\sigma(x') \in [v']_{w'}$.

So $C \in K_{w,w'}$ iff there is a substitution σ such that $w \xrightarrow{\sigma} w'$ and C maps $[v]_w$ to $[\sigma(v)]_{w'}$, for each $v \in \mathrm{Var}^*$ for which $[v]_w$ is non-empty.

We then have

$$(C \circ g_w)(v) = C(g_w(v)) = C([v]_w) = [\sigma(v)]_{w'} = g_{w'}(\sigma(v)) = (g_{w'} \circ \sigma)(v).$$

In sum, w' and C satisfy the requirement that wRw' and $C \in K_{w,w'}$ iff there

is a substitution σ such that $w \xrightarrow{\sigma} w'$ and $C \circ g_w = g_{w'} \circ \sigma$, So $\mathfrak{M}, w, g_w \Vdash \Box A$ iff $\mathfrak{M}, w', g_{w'} \circ \sigma \Vdash A$ for all w', σ such that $w \xrightarrow{\sigma} w'$.

Lemma 4.6 (Truth Lemma).
If $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ is the canonical model for a positive or negative logic L, $w \in W$, and g_w is the canonical assignment for w, then for any $\mathfrak{L}$-sentence A,

$$\mathfrak{M}, w, g_w \Vdash A \text{ iff } A \in w.$$

Proof. by induction on A.

1. A is $Px_1 \dots x_n$.

   By definition 2.9, $\mathfrak{M}, w, g_w \Vdash Px_1 \dots x_n$ iff $\langle g_w(x_1), \dots, g_w(x_n) \rangle \in I_w(P)$. By definition 4.5, $g_w(x_i)$ is $[x_i]_w$ or undefined if $[x_i]_w = \emptyset$. More-over, $I_w(P) = \{\langle [z_1]_w, \dots, [z_n]_w \rangle : Pz_1 \dots z_n \in w\}$ by definition 4.4 and, for the identity predicate, by the fact $I_w(=)$ is $\{\langle d, d \rangle : d \in U_w\} = \{\langle [z]_w, [z]_w \rangle : z = z \in w\} = \{\langle [z_1]_w, [z_2]_w \rangle : z_1 = z_2 \in w\}$.

   Now if $\langle g_w(x_1), \dots, g_w(x_n) \rangle \in I_w(P)$, then $\langle [x_1]_w, \dots, [x_n]_w \rangle \in \{\langle [z_1]_w, \dots, [z_n]_w \rangle : Pz_1 \dots z_n \in w\}$, where all the $[x_i]_w$ are non-empty (for $g_w(x_i)$ is defined). This means that there are variables $z_1, \dots, z_n$ such that $\{x_1 = z_1, \dots, x_n = z_n, Pz_1 \dots z_n\} \subseteq w$. Then $Px_1 \dots x_n \in w$ by (LL*).

   In the other direction, if $Px_1 \dots x_n \in w$ then $x_i = x_i \in w$ for all x_i in $x_1 \dots x_n$. Hence $\langle [x_1]_w, \dots, [x_n]_w \rangle \in \{\langle [z_1]_w, \dots, [z_n]_w \rangle : Pz_1 \dots z_n \in w\}$, and $\langle g_w(x_1), \dots, g_w(x_n) \rangle \in I_w(P)$.

2. A is $\neg B$.

   $\mathfrak{M}, w, g_w \Vdash \neg B$ iff $\mathfrak{M}, w, g_w \nVdash B$ by definition 2.9, iff $B \notin w$ by induction hypothesis, iff $\neg B \in w$ by definition 4.1.

3. A is $B \supset C$.

$\mathfrak{M}, w, g_w \Vdash B \supset C$ iff $\mathfrak{M}, w, g_w \nVdash B$ or $\mathfrak{M}, w, g_w \Vdash C$ by definition 2.9, iff $B \notin w$ or $C \in w$ by induction hypothesis, iff $B \supset C \in w$ by definition 4.1 and the fact that $B \supset C$ is L-entailed by $\neg B$ and by C.

4. *A* is $\forall x B$.

   By definition 2.9, $\mathfrak{M}, w, g_w \Vdash \forall x B$ iff $\mathfrak{M}, w, g_w{}^{x \mapsto d} \Vdash B$ for all $d \in D_w$. By definition 4.4, $D_w = \{[x]_w : x \in \mathrm{Var}^* \text{ and } E!x \in w\}$. So $\mathfrak{M}, w, g_w \Vdash \forall x B$ iff $\mathfrak{M}, w, g_w{}^{x \mapsto [y]_w} \Vdash B$ for all y with $E!y \in w$, iff $\mathfrak{M}, w, g_w \circ [y/x] \Vdash B$ for all y with $E!y \in w$, iff $\mathfrak{M}, w, g_w \Vdash [y/x]B$ for all y with $E!y \in w$ by lemma 3.1, iff $[y/x]B \in w$ for all y with $E!y \in w$ by induction hypothesis.

   It remains to show that $\forall x B \in w$ iff $[y/x]B \in w$ for all y with $E!y \in w$.

   LTR. Assume $\forall x B \in w$ and $E!y \in w$. By (FUI), then $[y/x]B \in w$.

   RTL. Assume $\forall x B \notin w$. Then $\exists x \neg B \in w$. Since w is witnessed, it contains $[y/x]\neg B$ and $E!y$ for some $y \notin \mathrm{FV}(B)$, and so w doesn't contain $[y/x]B$.

5. *A* is $\Box B$. $\mathfrak{M}, w, g_w \Vdash \Box B$ iff $w', g_{w'} \circ \sigma \Vdash B$ for all w', σ with $w \xrightarrow{\sigma} w'$ by lemma 4.5, iff $w', g_{w'} \Vdash \sigma(B)$ for all such w', σ by the substitution lemma 3.1, iff $\sigma(B) \in w'$ for all w', σ with $w \xrightarrow{\sigma} w'$ by induction hypothesis.

   It remains to show that

   $$\Box B \in w \text{ iff } \sigma(B) \in w' \text{ for all } w', \sigma \text{ with } w \xrightarrow{\sigma} w'.$$

   LTR. Assume $\Box B \in w$, and let w' and σ satisfy the condition that $\{\sigma(X) : \Box X \in w\} \subseteq w'$. Then $\sigma(B) \in w'$.

   RTL. Assume $\Box B \notin w$. Let σ be any injective substitution whose range excludes infinitely many variables. We show that $\{\sigma(X) : \Box X \in w\} \cup \{\neg\sigma(A)\}$ is L-consistent. Suppose not. Then there are $X_1..X_n$ such that $\Box X_i \in w$ and $\vdash_L \sigma(X_1) \wedge ... \wedge \sigma(X_n) \supset \sigma(A)$. Since the injective substitutions leave every formula unchanged, this means that $\vdash_L X_1 \wedge ... \wedge X_n \supset A$. By (K) and (Nec), it follows that $\vdash_L \Box X_1 \wedge ... \wedge \Box X_n \supset \Box A$.

Since all $\square X_i \in w$ and w is maximal consistent, it follows that $\square B \in w$. This contradicts our assumption that $\square B \notin w$.

So $\{\sigma(X) : \square X \in w\} \cup \{\neg\sigma(B)\}$ is L-consistent, and infinitely many variables don't occur in it. By the extensibility lemma 4.4, there is a Henkin set w' that extends $\{\sigma(X) : \square X \in w\} \cup \{\neg\sigma(B)\}$. We have $w \overset{\sigma}{\rightarrow} w'$, but $\sigma(B) \notin w'$.

4.3 Completeness of the base logics

We know from theorem 3.5 that FK is sound with respect to the class of total functional counterpart structures. In the previous section, we have essentially shown that FK is (strongly) complete with respect to this class. What we've shown is this:

Lemma 4.7 (Completeness lemma).
Every positive or negative modal predicate logic is strongly complete with respect to any class of counterpart structures that contains the structure of its canonical model.

Proof. Let L be a positive or negative modal predicate logic, and $\mathfrak{M}$ its canonical model. Assume some set Γ of $\mathfrak{L}$-formulas is L-consistent. By the extensibility lemma 4.4, Γ is contained in some Henkin set w for L. (Note that Γ contains no variables from Var^*.) By the truth lemma 4.6, $\mathfrak{M}, w, g_w \Vdash A$ for each $A \in \Gamma$. So Γ is satisfiable any class of structures that contains the structure of $\mathfrak{M}$.

Lemma 4.8 (Canonicity of FK).
The structure of the canonical model for FK is total and functional.

Proof. Immediate from lemmas 4.3 and 4.2.

Theorem 4.9 (Completeness of FK**).**
The system FK *is strongly complete with respect to the class of total functional counterpart structures.*

Proof. Immediate from lemmas 4.7 and 4.8.

As for QK, we know from theorem 3.7 that QK is sound with respect to the class of classical counterpart structures. To show completeness, we only need to show that the structure of the canonical model for QK is classical.

Lemma 4.10 (Canonicity of QK**).**
The structure of the canonical model for QK *is total, single-domain, and functional.*

Proof. Immediate from lemmas 4.3 and 4.2, given the fact that QK is both positive and negative. (For an alternative proof that the canonical model for QK is single-domain, note that $E!t$ is provable in classical logic from the (UI) instance $\forall x \; x \neq t \supset t \neq t$ and the (SI) instance $t = t$. So every world w in the canonical model for QK contains $E!t$, for every term t. And so $D_w = U_w$, by definition 4.4.)

Theorem 4.11 (Completeness of QK**).**
The system QK *is strongly complete with respect to the class of classical counterpart structures.*

Proof. Immediate from lemmas 4.7 and 4.10.

As for NK, we know from theorem 3.9 that NK is sound with respect to the class of single-domain functional counterpart structures. To show completeness, we need to show that the structure of the canonical model for NK is functional and single-domain.

Lemma 4.12 (Canonicity of NK**).**
The structure of the canonical model for NK *is single-domain and functional.*

Proof. Immediate from lemmas 4.3 and 4.2.

Theorem 4.13 (Completeness of NK**).**
The system NK *is strongly complete with respect to the class of single-domain functional counterpart structures.*

Proof. Immediate from lemmas 4.7 and 4.12.

These simple results are noteworthy because they cannot be obtained with any form of Kripke semantics. FK, QK, and NK are *incomplete* in Kripke semantics, meaning that there is no class of Kripke structures with respect to which they are sound and complete. An easy way to see this is to note that the "necessity of distinctness"

(ND) $x \neq y \supset \Box x \neq y$

is valid in Kripke semantics, but not provable in any of our three base logics.

In chapter 6, we are going to explore the completeness of systems that extend our base logics by adding familiar axioms such as (T), (4), or (BF). In each case, we know from lemma 4.7 that the extended system is strongly complete with respect to any class of counterpart structures that contains the structure of its canonical model. It remains to show that the system is sound with respect to this structure. To this end, it will be useful to have a general idea of what a structure for the relevant system must look like. This is the topic of the next chapter.

5 Correspondence

5.1 Schemas, frames, and properties of the image relation

A well-known feature of Kripke semantics for propositional modal logic is that various modal schemas correspond to conditions on the accessibility relation, in the sense that the schema is valid on all and only those Kripke frames whose accessibility relation satisfies the condition: $\Box A \supset A$ *corresponds to* (or *defines*) reflexivity, $A \supset \Box \Diamond A$ corresponds to symmetry, and so on.

In counterpart semantics, these same schemas constrain both the accessibility and the counterpart relations. Consider $\Box Fx \supset Fx$. Loosely speaking, the antecedent $\Box Fx$ is true at w iff all counterparts of x are F at all accessible worlds. This does not entail that x is F at w unless (i) w can see itself *and* (ii) x is its own counterpart at w.

Before we generalize this observation, let's first review some familiar definitions from propositional modal logic.

> **Definition 5.1 (Languages of propositional modal logic).**
>
> A set of formulas $\mathfrak{L}_0$ is a *(unimodal) propositional language* if there is a denumerable set of expressions *Prop* (the *sentence letters* of $\mathfrak{L}_0$) distinct from $\{\neg, \supset, \Box\}$ such that $\mathfrak{L}_0$ is generated by the rule
>
> $$P \mid \neg A \mid (A \supset B) \mid \Box A,$$
>
> where $P \in Prop$.

Note that the language $\mathfrak{L}$ of modal predicate logic is a unimodal propositional language if we define a "sentence letter" as any $\mathfrak{L}$-formula that isn't of the form

$\neg A, A \supset B$ or $\Box A$. (On this usage, $\forall x \Box (Fx \supset Gx)$, for example, is a sentence letter.) Let's call such $\mathfrak{L}$-formulas *quasi-atomic*.

Definition 5.2 (Frames and valuations).

A *frame* is a pair consisting of a non-empty set W and a relation $R \subseteq W^2$.

A *valuation of* a unimodal propositional language $\mathfrak{L}_0$ *on* a frame $\mathfrak{F} = \langle W, R \rangle$ is a function V that maps every sentence letter of $\mathfrak{L}_0$ to a subset of W.

Definition 5.3 (Propositional truth).

For any frame $\mathfrak{F} = \langle W, R \rangle$, point $w \in W$, valuation V on $\mathfrak{F}$, sentence letter P and $\mathfrak{L}_0$-sentences A and B,

$\mathfrak{F}, V, w \Vdash_0 P$ iff $w \in V(P)$,

$\mathfrak{F}, V, w \Vdash_0 \neg A$ iff $\mathfrak{F}, V, w \nVdash_0 A$,

$\mathfrak{F}, V, w \Vdash_0 A \supset B$ iff $\mathfrak{F}, V, w \nVdash_0 A$ or $\mathfrak{F}, V, w \Vdash_0 B$,

$\mathfrak{F}, V, w \Vdash_0 \Box A$ iff $\mathfrak{F}, V, w' \Vdash_0 A$ for all w' with wRw'.

Definition 5.4 (Frame validity).

A formula A of a unimodal propositional language is *valid on* a frame $\mathfrak{F} = \langle W, R \rangle$ if $\mathfrak{F}, V, w \Vdash_0 A$ for all $w \in W$ and valuations V of the language on $\mathfrak{F}$.

Definition 5.5 (Frame correspondence).

A sentence A of a unimodal propositional language *defines* or *corresponds to* a class of frames $\mathfrak{C}$ iff $\mathfrak{C}$ is the class of frames on which A is valid.

In propositional modal logic, formulas are true or false relative to a world; the box shifts the world of evaluation. In counterpart semantics, formulas are true or false relative to a pair w, g of a world and an assignment function, and the box shifts these points of evaluation: $\Box A$ is true at w, g iff A is true at all w', g' such that $w, g \rhd w', g'$. This suggests the following *conjecture*: if a schema of propositional modal logic corresponds to a certain property of the accessibility relation R of Kripke frames,

then it corresponds to the same property of the relation $\rhd$ in counterpart structures. That is, if a schema A is valid on all and only the Kripke frames whose accessibility relation satisfies a certain condition, then the schema is valid on all and only the counterpart structures whose image relation (on w, g pairs) satisfies this condition.

It is, in fact, easy to show that the schema will be valid on *all* those counterpart structures. But I am not able to show that it is valid on *only* those structures. The problem is that any $\mathfrak{L}$-instance of a propositional schema A only contains finitely many free variables. A formula of $\mathfrak{L}$ is true or false only relative to a world w and a *finite fragment* of an assignment g on U_w – a fragment that interprets the free variables in the formula.

> **Open question:** Give an example of a schema that is valid on all and only the propositional frames in which R has a certain property, but that is not valid on all and only the counterpart structures in which $\rhd$ has that property. Alternatively, prove that there is no such schema.

5.2 Finitary satisfaction

In definition 2.9, the "image" relation $\rhd$ between world-assignment pairs plays the role of the accessibility relation in standard Kripke semantics. We could have taken $\rhd$ as primitive, instead of deriving it from K and R. Conceptually, this would have the downside of blurring the distinction between the structures on which our language is interpreted and the interpretation. Assignments belong to the interpretation; they don't represent an independent aspect of a scenario in which $\mathfrak{L}$-formulas are interpreted. From a technical point of view, taking $\rhd$ as primitive would also require formulating some constraints to ensure, for example, that if $w, g \rhd w', g'$ then $w, g^{x \mapsto g(y)} \rhd w', g'^{x \mapsto g'(y)}$.

That said, it can be useful to think of the $\rhd$ relation as an extended counterpart relation or accessibility relation that relates *finite sequences of individuals at worlds*.

Let's recast definition 2.9 in terms of this relation. To keep things simple, I will assume that assignments and structures are total, so that we can ignore gappy sequences that would arise in a negative semantics.

> **Definition 5.6 (n-sequential accessibility).**
> Given a counterpart structure $\mathfrak{S} = \langle W, R, U, D, K \rangle$ and a number $n \in \mathbb{N}$, the *n-sequential accessibility relation of $\mathfrak{S}$* is the (smallest) relation R^n that holds between an $n+1$-tuple $\langle w, d_1, \ldots, d_n \rangle$ and an $n+1$-tuple $\langle w', d'_1, \ldots, d'_n \rangle$ iff there is a $C \in K_{w,w'}$ for which $d_1 C d'_1, \ldots, d_n C d'_n$.

(Note that $R^0 = R$.)

We are going to evaluate formulas relative to sequences $\langle w, d_1, \ldots, d_n \rangle$ consisting of a world and some individuals: the alphabetically first term will pick out d_1, the second d_2, and so on. A formula can only be evaluated if the sequence provides a value for each of its term.

> **Definition 5.7 (Rank).**
> Let ρ be some fixed "alphabetical" order on the terms T in a language, i.e. a bijection from T into $\mathbb{N}^+$. I will use 'v' for the inverse of ρ, so that v_1 is the alphabetically first term, v_2 the second, and so on. The *rank* of an $\mathfrak{L}$-formula A is the smallest number $r \in \mathbb{N}$ such that all members of $\text{Var}(A)$ have a ρ-value less than or equal to r.

> **Definition 5.8 (Finitary satisfaction).**
> Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ be a counterpart model, w a member of W, and $d_1, \ldots, d_r$ (not necessarily distinct) elements of U_w. For any number $n \in \mathbb{N}$ and $\mathfrak{L}$-formulas $Pt_1 \ldots t_m, A$, and B whose rank is less than or equal to n, define
>
> $\mathfrak{M}, w, d_1, \ldots, d_n \Vdash Pt_1 \ldots t_m$ iff $\langle d_{\rho(t_1)}, \ldots, d_{\rho(t_m)} \rangle \in I_w(P)$.
>
> $\mathfrak{M}, w, d_1, \ldots, d_n \Vdash \neg A$ iff $\mathfrak{M}, w, d_1, \ldots, d_n \nVdash A$.
>
> $\mathfrak{M}, w, d_1, \ldots, d_n \Vdash A \supset B$ iff $\mathfrak{M}, w, d_1, \ldots, d_n \nVdash A$ or $\mathfrak{M}, w, d_1, \ldots, d_n \Vdash B$.
>
> $\mathfrak{M}, w, d_1, \ldots, d_n \Vdash \forall x A$ iff $\mathfrak{M}, w, d_1', \ldots, d_n' \Vdash A$ for all $d_1', \ldots, d_n'$ such that $d_{\rho(x)}' \in D_w$ and $d_i' = d_i$ for all $i \neq \rho(x)$.

$$\mathfrak{M}, w, d_1, \dots, d_n \Vdash \Box A \qquad \text{iff } w', d_1{}', \dots, d_n{}' \Vdash A \text{ for all } w, d_1{}', \dots, d_n{}'$$
$$\text{such that } \langle w, d_1, \dots, d_n \rangle R^n \langle w', d_1{}', \dots, d_n{}' \rangle,$$

where R^n is the *n*-sequential counterpart relation of $\mathfrak{S}$.

Lemma 5.1.
For any total counterpart model $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$, world $w \in W$, (total) assignment g on U_w, formula A, and number $n \geq \rho(A)$,

$$\mathfrak{M}, w, g \Vdash A \text{ iff } \mathfrak{M}, w, g(v_1), \dots, g(v_n) \Vdash A.$$

Proof. by induction on A.

- A is $Pt_1 \dots t_m$. $\mathfrak{M}, w, g \Vdash Pt_1 \dots t_m$ iff $\langle g(t_1), \dots, g(t_m) \rangle \in I_w(P)$ by definition 2.9, iff $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \Vdash Px_1 \dots x_m$ by definition 5.8. For the last step, note that the *i*th element of $\langle g(v_1), \dots, g(v_n) \rangle$ is $g(v_i)$; so if t_i is v_j then the $\rho(t_i)$-th element of $\langle g(v_1), \dots, g(v_n) \rangle$ is $g(v_j)$, which is $g(t_i)$.

- A is $\neg B$. $\mathfrak{M}, w, g \Vdash \neg B$ iff $\mathfrak{M}, w, g \nVdash B$ by definition 2.9, iff $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \nVdash B$ by induction hypothesis, iff $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \Vdash \neg B$ by definition 5.8.

- A is $B \supset C$. $\mathfrak{M}, w, g \Vdash B \supset C$ iff $\mathfrak{M}, w, g \nVdash B$ or $\mathfrak{M}, w, g \Vdash C$ by definition 2.9, iff $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \nVdash B$ or $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \Vdash C$ by induction hypothesis, iff $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \Vdash B \supset C$ by definition 5.8.

- A is $\forall x B$. $\mathfrak{M}, w, g \Vdash \forall x B$ iff $\mathfrak{M}, w, g^{x \mapsto d} \Vdash B$ for all $d \in D_w$ by definition 2.9, iff $\mathfrak{M}, w, g^{x \mapsto d}(v_1), \dots, g^{x \mapsto d}(v_n) \Vdash B$ for all $d \in D_w$ by induction hypothesis, iff $\mathfrak{M}, w, g(v_1)', \dots, g(v_n)' \Vdash B$ for all $g(v_1)', \dots, g(v_n)'$ such that $g(v_{\rho(x)})' \in D_w$ and $g(v_i)' = g(v_i)$ for all $i \neq \rho(x)$, iff $\mathfrak{M}, w, g(v_1), \dots, g(v_n) \Vdash \forall x B$ by definition 5.8.

- A is $\Box B$. By definition 2.9, $\mathfrak{M}, w, g \Vdash \Box B$ iff $\mathfrak{M}, w', g' \Vdash B$ for all w', g' such that $w, g \rhd w', g'$. By induction hypothesis, it follows that $\mathfrak{M}, w, g \Vdash \Box B$ iff

$$\mathfrak{M}, w', g'(v_1), \dots, g'(v_n) \Vdash B \text{ for all } w', g' \text{ such that } w, g \rhd w', g'. \tag{1}$$

63

From the other direction, $\mathfrak{M}, w, g(v_1), \ldots, g(v_n) \Vdash \Box B$ iff

$$\mathfrak{M}, w', d'_1, \ldots, d'_n \Vdash B \text{ for all } w', d'_1, \ldots, d'_n$$
$$\text{s.t. } w, g(v_1), \ldots, g(v_n) R^n w', d'_1, \ldots, d'_n. \quad (2)$$

It remains to show that (1) and (2) are equivalent.

For the forward direction, let $w', d'_1, \ldots, d'_n$ be such that $w, g(v_1), \ldots, g(v_n) R^n$ $w', d'_1, \ldots, d'_n$. By definition 5.6, this means that there is a $C \in K_{w,w'}$ such that $g(v_i) C d'_i$ for all $i \in \{1..n\}$. Let g' be any assignment with

$$g'(v_i) = \begin{cases} d'_i & \text{if } i \in \{1..n\} \\ \text{an arbitrary } d' \text{ with } g(v_i) C d' & \text{otherwise.} \end{cases}$$

There must be some such g' because C it total. So $\mathfrak{M}, w', g'(v_1), \ldots, g'(v_n) \Vdash B$ by (1). And so $\mathfrak{M}, w', d'_1, \ldots, d'_n \Vdash B$ by construction of g'. So we can derive (2) from (1).

For the backward direction, let w', g' be such that $w, g \rhd w', g'$. By definition 2.8, this means that there is a $C \in K_{w,w'}$ such that $g(v_i) C g'(v_i)$ for all variables v_i. So $w, g(v_1), \ldots, g(v_n) R^n w', g'(v_1), \ldots, g'(v_n)$, by definition 5.6. By (2), it follows that $\mathfrak{M}, w', g'(v_1), \ldots, g'(v_n) \Vdash B$.

5.3 Propositional guises

We've seen that the semantics from section 2.5 can be expressed in terms of a relation between formulas on the one hand and sequences of a world and finitely many individuals on the other. If we re-interpret these sequences as "worlds", and treat quasi-atomic sentences as atomic, the semantics reduces to standard Kripke semantics for propositional modal logic.

Definition 5.9 (N-ary Opaque Propositional Guise).
The *n-ary opaque propositional guise of a counterpart structure* $\mathfrak{S} = \langle W, R, U, D, K \rangle$

is the Kripke frame $\langle W^n, R^n \rangle$ where W^n is the set of points $\langle w, d_1, \dots, d_n \rangle$ such that $w \in W$, $d_1 \in U_w, \dots, d_n \in U_w$, and R^n is the n-ary accessibility relation for $\mathfrak{S}$.

The *n-ary opaque propositional guise of a predicate interpretation I* on $\mathfrak{S}$ is the valuation function V^n on $\langle W^n, R^n \rangle$ such that for every quasi-atomic formula $A \in \mathfrak{L}$, $V^n(A) = \{\langle w, d_1, \dots, d_n \rangle : \mathfrak{S}, I, w, d_1, \dots, d_n \Vdash A\}$.

Lemma 5.2 (Truth-preservation under opaque guises).
For any total counterpart structure $\mathfrak{S} = \langle W, R, U, D, K \rangle$, interpretation I on $\mathfrak{S}$, world $w \in W$, individuals $d_1, \dots, d_n \in U_w$, and $\mathfrak{L}$-formula A with rank $\leq n$,

$$\mathfrak{S}, I, w, d_1, \dots, d_n \Vdash A \text{ iff } \mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \Vdash_0 A,$$

where $\mathfrak{S}^n$ and V^n are the n-ary opaque propositional guises of $\mathfrak{S}$ and I respectively.

Proof. We argue by induction on A, where quasi-atomic formulas all have complexity zero.

- A is quasi-atomic. By definition 5.9, $V^n(A) = \{\langle w, d_1, \dots, d_n \rangle : \mathfrak{S}, I, w, d_1, \dots, d_n \Vdash A\}$. So $\mathfrak{S}, I, w, d_1, \dots, d_n \Vdash A$ iff $\langle w, d_1, \dots, d_n \rangle \in V^n(A)$, iff $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \Vdash_0 A$ by definition 5.3.

- A is $\neg B$. $\mathfrak{S}, I, w, d_1, \dots, d_n \Vdash \neg B$ iff $\mathfrak{S}, I, w, d_1, \dots, d_n \nVdash B$ by definition 5.8, iff $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \nVdash_0 B$ by induction hypothesis, iff $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \Vdash_0 \neg B$ by definition 5.3.

- A is $B \supset C$. $\mathfrak{S}, I, w, d_1, \dots, d_n \Vdash B \supset C$ iff $\mathfrak{S}, I, w, d_1, \dots, d_n \nVdash B$ or $\mathfrak{S}, I, w, d_1, \dots, d_n \Vdash C$ by definition 5.8, iff $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \nVdash_0 B$ or $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \Vdash_0 C$ by induction hypothesis, iff $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \Vdash_0 B \supset C$ by definition 5.3.

- A is $\square B$. $\mathfrak{S}, I, w, d_1, \dots, d_n \Vdash \square B$ iff $\mathfrak{S}, I, w', d_1', \dots, d_n' \Vdash B$ for all $w, d_1', \dots, d_r'$ such that $\langle w, d_1, \dots, d_r \rangle R^n \langle w', d_1', \dots, d_n' \rangle$ by definition 5.8, iff $\mathfrak{S}^n, V^n, \langle w', d_1', \dots, d_n' \rangle \Vdash_0$

B for all such $w, d_1', \dots, d_r'$ by induction hypothesis, iff $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \Vdash_0$ $\Box B$ by definition 5.3.

Lemma 5.3 (Finite correspondence transfer).

If A is a formula of (unimodal) propositional modal logic that is valid on all and only the Kripke frames in some class $\mathfrak{C}$, and $n \in \mathbb{N}$, then the $\mathfrak{L}$-formulas that result from A by uniformly substituting sentence letters by $\mathfrak{L}$-formulas of rank $\leq n$ are (jointly) valid on a total counterpart structure $\mathfrak{S} = \langle W, R, U, D, K \rangle$ iff the n-ary opaque propositional guise of $\mathfrak{S}$ is in $\mathfrak{C}$.

Proof. Assume A is valid on all and only the Kripke frames in $\mathfrak{C}$, and let $p_1, \dots, p_k$ be the sentence letters in A. Let $\mathfrak{S} = \langle W, R, U, D, K \rangle$ be a total counterpart structure whose n-ary opaque propositional guise $\langle W^n, R^n \rangle$ is in $\mathfrak{C}$.

Suppose for reductio that some formula A' is not valid on $\mathfrak{S}$, and A' results from A by uniformly substituting the sentence letters p_i in A by $\mathfrak{L}$-formulas $p_i^{\mathfrak{L}}$ of rank $\leq n$. Then there is an interpretation I on $\mathfrak{S}$, a world $w \in W$, and an assignment g on U_w such that $\mathfrak{S}, I, w, g \not\Vdash A'$. By lemma 5.1, this means that $w, d_1, \dots, d_r \not\Vdash A'$, where $d_1 = g(v_1), \dots, d_r = g(v_r)$. By lemma 5.2, it follows that $\mathfrak{S}^n, V^n, \langle w, d_1, \dots, d_n \rangle \not\Vdash_0 A'$. But then $\mathfrak{S}^n, V^{n'}, \langle w, d_1, \dots, d_n \rangle \not\Vdash_0 A$, where $V^{n'}$ is such that for all sentence letters p_i in A, $V^{n'}(p_i) = V^n(p_i^{\mathfrak{L}})$. This contradicts the assumption that A is valid on $\langle W^n, R^n \rangle$.

We also have to show that the relevant $\mathfrak{L}$-formulas are valid *only* on structures $\mathfrak{S}$ whose n-ary opaque propositional guise is in $\mathfrak{C}$. So let $\mathfrak{S}$ be a structure whose guise $\langle W^n, R^n \rangle$ is not in $\mathfrak{C}$. Since A is valid only on frames in $\mathfrak{C}$, we know that there is some valuation V on $\langle W^n, R^n \rangle$ and some $\langle w, d_1, \dots, d_n \rangle \in W^n$ such that $W^n, R^n, V, \langle w, d_1, \dots, d_n \rangle \not\Vdash_0 A$. Let A' result from A by uniformly substituting each sentence letter p_i in A by an n-ary predicate P_i followed by the variables $v_1 \dots v_n$, with distinct predicates for distinct sentence letters. Let I be a predicate interpretation such that for all P_i and $w' \in W$, $I_{w'}(P_i) = \{\langle d_1', \dots, d_n' \rangle : \langle w', d_1', \dots, d_n' \rangle \in V(p_i)\}$. A simple induction on subformulas B of A shows

that for all $\langle w', d'_1, \ldots, d'_n \rangle \in W^n$, we have $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \Vdash_0 B$ iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \Vdash B'$, where B' is B with all p_i replaced by $P_i v_1 \ldots v_n$. Given that $W^n, R^n, V, \langle w, d_1, \ldots, d_n \rangle \nVdash_0 A$, it follows that $\mathfrak{S}, I, w, d_1, \ldots, d_n \nVdash A'$.

Here is the simple induction.

1. B is a sentence letter p_i. Then B' is $P_i v_1 \ldots v_n$. We have W^n, R^n, V, $\langle w', d'_1, \ldots, d'_n \rangle \Vdash_0 p_i$ iff $\langle w', d'_1, \ldots, d'_n \rangle \in V(p_i)$ by definition 5.3, iff $\langle d'_1, \ldots, d'_n \rangle \in I_w(P_i)$ by construction of I, iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \Vdash P_i v_1 \ldots v_n$ by definition 5.8.

2. B is $\neg C$. Then B' is $\neg C'$, where C' is C with all p_i replaced by $P_i v_1 \ldots v_n$. We have $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \Vdash_0 \neg C$ iff $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \nVdash_0 C$ by definition 5.3, iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \nVdash C'$ by induction hypothesis, iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \Vdash \neg C'$ by definition 5.8.

3. B is $C \supset D$. Then B' is $C' \supset D'$, where C' and D' are C and D respectively with all p_i replaced by $P_i v_1 \ldots v_n$. We have $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \Vdash_0 C \supset D$ iff $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \nVdash_0 C$ or $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \Vdash_0 D$ by definition 5.3, iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \nVdash C'$ or $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \Vdash D'$ by induction hypothesis, iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \Vdash C' \supset D'$ by definition 5.8.

4. B is $\Box C$. Then B' is $\Box C'$, where C' is C with all p_i replaced by $P_i v_1 \ldots v_n$. We have $W^n, R^n, V, \langle w', d'_1, \ldots, d'_n \rangle \Vdash_0 \Box C$ iff $W^n, R^n, V, \langle w'', d''_1, \ldots, d''_n \rangle \Vdash_0 C$ for all $\langle w'', d''_1, \ldots, d''_n \rangle$ with $\langle w', d'_1, \ldots, d'_n \rangle R^n \langle w'', d''_1, \ldots, d''_n \rangle$ by definition 5.3, iff $\mathfrak{S}, I, w'', d''_1, \ldots, d''_n \nVdash C'$ for all such $\langle w'', d''_1, \ldots, d''_n \rangle$ by induction hypothesis, iff $\mathfrak{S}, I, w', d'_1, \ldots, d'_n \Vdash \Box C'$ by definition 5.8.

Given that a modal schema restricted to the variables $v_1, \ldots, v_n$ defines a constraint on the n-sequential accessibility relation of a counterpart structure $\mathfrak{S}$, the unrestricted schema defines a constraint on all sequential accessibility relations. Let's fold these into a single entity.

Definition 5.10 (Sequential accessibility relation).
The *sequential accessibility relation* R^* of a total counterpart structure $\mathfrak{S}$ is the union of the *n*-sequential accessibility relations R^n of $\mathfrak{S}$. That is, $R^* = \bigcup_{n \in \mathbb{N}} R^n$.

Definition 5.11 (Opaque Propositional Guise).
The *opaque propositional guise* of a counterpart structure $\mathfrak{S} = \langle W, R, U, D, K \rangle$ is the disjoint union of the *n*-ary opaque propositional guises of $\mathfrak{S}$, i.e. the Kripke frame $\langle W^*, R^* \rangle$ such that R^* is the sequential accessibility relation of $\mathfrak{S}$ and W^* is the set of points w^* such that for some $n \in \mathbb{N}$, world $w \in W$ and individuals $d_1, \ldots, d_n \in U_w$, $w^* = \langle w, d_1, \ldots, d_n \rangle$.

Theorem 5.4 ((Positive) correspondence transfer).
If A is a formula of (unimodal) propositional modal logic that is valid on all and only the Kripke frames in some class $\mathfrak{C}$, then the $\mathfrak{L}$-formulas that result from A by uniformly substituting sentence letters by $\mathfrak{L}$-formulas are (jointly) valid on a total counterpart structure $\mathfrak{S}$ iff the opaque propositional guise of $\mathfrak{S}$ is in $\mathfrak{C}$.

Proof. Since validity in propositional modal logic is preserved under disjoint unions, *A* is valid on the opaque propositional guise of a structure $\mathfrak{S}$ iff *A* is valid on each *n*-ary opaque propositional guise of $\mathfrak{S}$, with $n \in \mathbb{N}$. (See, e.g., [Blackburn et al. 2001], p.140, Theorem 3.14.(i).) So the opaque propositional guise of $\mathfrak{S}$ is in $\mathfrak{C}$ iff all *n*-ary opaque propositional guises of $\mathfrak{S}$ are in $\mathfrak{C}$.

Assume *A* is valid on all and only the Kripke frames in $\mathfrak{C}$. Let A' be an $\mathfrak{L}$-formula that results from *A* by uniformly substituting sentence letters by $\mathfrak{L}$-formulas. By lemma 5.3, A' is valid on all total counterpart structures $\mathfrak{S}$ whose *n*-ary propositional guise is in $\mathfrak{C}$, where *n* is the rank of A'. Any total structure whose propositional guise is in $\mathfrak{C}$ satisfies this condition.

To show that the *A* schema is valid *only* on structures $\mathfrak{S}$ whose guise is in $\mathfrak{C}$, let $\mathfrak{S}$ be a structure whose guise is not in $\mathfrak{C}$. Then there is some *n* such that the

n-ary guise of $\mathfrak{S}$ is not in $\mathfrak{C}$. By lemma 5.3, there is an $\mathfrak{L}$-substitution instance A' of A with rank *n* that is not valid on $\mathfrak{S}$.

As a union of relations of different arity, R^* is a somewhat gerrymandered entity. It may help to understand statements about R^* as universal statements about its members R^n. For example, the schema $\Box A \supset A$ is valid iff (0) every world can see itself, and (1) every individual at every world is its own counterpart (relative to some counterpart relation), and (2) every pair of individuals at every world is its own counterpart (relative to some counterpart relation), and so on.

(Why "relative to *some* counterpart relation"? Because definition 5.6 says that w, dR^1w', d' iff there is a $C \in K_{w,w'}$ such that dCd'. So w, dR^1w, d iff there is a $C \in K_{w,w}$ such that dCd.)

5.4 Interaction principles

So far, we have focussed on propositional schemas. We may also ask what kinds of structures are defined by axioms or schemas that make use of the first-order machinery in $\mathfrak{L}$, like these:

(BF) $\forall x \Box A \supset \Box \forall x A$

(CBF) $\Box \forall x A \supset \forall x \Box A$

(NE) $E!x \supset \Box E!x$

(NF) $\neg E!x \supset \Box \neg E!x$

(ND) $x \neq y \supset \Box x \neq y$

Most of the following facts are (in essence) shown in [Kutz 2000: 79ff.].

> **Theorem 5.5.**
> - (BF) *is valid on a counterpart structure iff the structure is D*-surjective, *meaning that if $C \in K_{w,w'}$ then every $d' \in D_{w'}$ is such that dCd' for some $d \in D_w$.*
> - (CBF) *and* (NE) *are valid on a counterpart structure iff the structure is* existence-preserving, *meaning that if $d \in D_w$ and dCd' for some*

$C \in K_{w,w'}$ *then* $d' \in D_{w'}$.
- (NF) *is valid on a counterpart structure iff the structure is* nonexistence-preserving, *meaning that if* $d \notin D_w$ *and* dCd' *for some* $C \in K_{w,w'}$ *then* $d' \notin D_{w'}$.
- (ND) *is valid on a counterpart structure iff the structure is* injective, *meaning that no two individuals have the same counterpart at any world, relative to a single counterpart relation.*

I should give proof of these, and possibly other relevant facts.

6 Completeness of stronger systems

6.1 Beginnings

In section 4.3, we showed completeness for the base logics FK, NK, and QK. In this chapter, we look at stronger systems that extend these base logics by further axioms.

> We'll see that the simple approach from chapter 4 doesn't work for many stronger systems. I have not yet found a general approach that works for, say, QS4M, QS4.4, and QS4.2 + BF. These are incomplete in Kripke semantics. I do not know whether they are complete in counterpart semantics (as presented in chapter 2).

Suppose, for example, that we extend FK by the (T) schema

(T) $\Box A \supset A$.

In standard Kripke semantics for propositional modal logic, the (T) schema is valid on all and only the reflexive frames. By lemma 5.4, it follows that the schema is valid on all and only the "locally reflexive" counterpart structures in which

$$w, d_1, \ldots, d_n R^n w, d_1, \ldots, d_n$$

for every $w \in W$, $n \in \mathbb{N}$, and $d_1, \ldots, d_n \in U_w$. By definition 5.6, this means that wRw and there is a $C \in K_{w,w}$ such that $d_1 C d_1, \ldots, d_n C d_n$.

Since the rules (closure conditions) of FK preserve validity on any frame (by lemma 3.4), we know that FK+(T) is sound with respect to the class of locally reflexive structures. We can establish completeness by showing that the canonical model for FK+(T) is locally reflexive.

Lemma 6.1.
If a logic L contains all instances of (T) *then the canonical model for L is locally reflexive.*

Proof. Let w be a world in the canonical model for L. Since w contains all instances of (T), we have $\{A : \Box A \in w\} \subseteq w$. So $w \overset{\text{id}}{\longrightarrow} w$, where id is the identity substitution that maps every variable to itself. (The $\overset{\sigma}{\longrightarrow}$ relation is defined in definition 4.3.) By definition 4.4, it follows that there is a $C \in K_{w,w}$ that maps every variable class $[x]_w$ to itself.

Proposition 6.2.
FK+(T) *is strongly complete with respect to the class of total functional structures that are locally reflexive.*

Proof. Immediate from lemmas 4.7, 4.8, and 6.1.

The same reasoning shows that QK+(T) is complete with respect to the class of locally reflexive classical structures.

Next, let's add the (4) schema

(4) $\Box A \supset \Box\Box A$.

This is valid on all and only the "locally transitive" counterpart structures, in which

$$w, d_1, \ldots, d_n R^n w', d'_1, \ldots, d'_n \text{ and } w', d'_1, \ldots, d'_n R^n w'', d''_1, \ldots, d''_n$$
$$\text{implies } w, d_1, \ldots, d_n R^n w'', d''_1, \ldots, d''_n.$$

We can establish completeness of (positive) systems containing (4) by showing that their canonical model is locally transitive.

Lemma 6.3.
If a logic contains all instances of (=R) *and* (4) *then its canonical model is locally transitive.*

Proof. Let w, w', w'' be worlds in the canonical model for a system L that contains all instances of schema (4), and suppose that

$$w, d_1, \ldots, d_n R^n w', d'_1, \ldots, d'_n \text{ and} \tag{1}$$
$$w', d'_1, \ldots, d'_n R^n w'', d''_1, \ldots, d''_n, \tag{2}$$

where the d_i are in U_w, the d'_i are in $U_{w'}$, and the d''_i are in $U_{w''}$. By definitions 5.6 and 4.4, (1) means that there is a substitution σ such that

(a) if $\Box A \in w$ then $\sigma(A) \in w'$, and

(b) for each i there is an $x \in d_i$ with $\sigma(x) \in d'_i$.

Similarly, (2) means that there is a substitution σ' such that

(c) if $\Box A \in w'$ then $\sigma'(A) \in w''$, and

(d) for each i there is a $y \in d'_i$ with $\sigma'(y) \in d''_i$.

Let $\sigma'' = \sigma' \circ \sigma$. We show that

(e) if $\Box A \in w$ then $\sigma''(A) \in w''$, and

(f) for each i there is an $x \in d_i$ with $\sigma''(x) \in d''_i$.

For (e), suppose $\Box A \in w$. Then $\Box\Box A \in w$ by (4), and so $\sigma''(A) \in w''$ by (a) and (c).

Now for (f). We know from (b) that there is an $x \in d_i$ with $\sigma(x) \in d'_i$. We also know from (d) that there is a $y \in d'_i$ with $\sigma'(y) \in d''_i$. By construction of d'_i and w', it follows that $\sigma(x) = y \in w'$. So $\Box\sigma(x) = y \in w'$ by lemma 3.14, and so $\sigma'(\sigma(x)) = \sigma'(y) \in w'''$ by (c).

Proposition 6.4.
- FK+(4) *is complete with respect to the class of total functional structures that are locally transitive.*

- QK+(4) *is complete with respect to the class of classical structures that are locally transitive.*
- FK+(T)+(4) *is complete with respect to the class of total functional structures that are locally reflexive and locally transitive.*
- QK+(T)+(4) *is complete with respect to the class of classical structures that are locally reflexive and locally transitive.*

Proof. Immediate from lemmas 4.7, 4.8, 4.10, 6.1, and 6.3.

So far, so good. Unfortunately, the approach we have taken breaks down for other prominent axioms, such as

(B) $A \supset \Box\Diamond A$,

(5) $\Diamond A \supset \Box\Diamond A$,

(M) $\Box\Diamond A \supset \Diamond\Box A$,

(Triv) $A \leftrightarrow \Box A$, or

(BF) $\forall x \Box A \supset \Box \forall x A$.

6.2 Restricted substitution models

Consider FK+(B). We'd like to show that the canonical model for FK+(B) is locally symmetric, so that

$$\text{if } w, d_1, \ldots, d_n R^n w', d_1', \ldots, d_n' \text{ then } w', d_1', \ldots, d_n' R^n w, d_1, \ldots, d_n.$$

Suppose $w, d_1, \ldots, d_n R^n w', d_1', \ldots, d_n'$. This means that there is a substitution σ such that

(a) if $\Box A \in w$ then $\sigma(A) \in w'$, and

(b) for each i there is an $x \in d_i$ with $\sigma(x) \in d_i'$.

We'd like to infer that there is a substitution σ' such that

(c) if $\Box A \in w'$ then $\sigma'(A) \in w$, and

(d) for each i there is a $y \in d'_i$ with $\sigma'(y) \in d_i$.

There is no guarantee that this is the case, since σ may not be invertible.

It turns out that we can restrict the eligible substitutions in definition 4.4 to bijective (and hence invertible) substitutions. We will make use of this in chapter 9. But this kind of move doesn't generalize.

For a simple illustration, consider (Triv). It is easy to see that (Triv) s not valid on any structure in which some world can see another: we can then falsify the simple (Triv) instance $P \leftrightarrow \Box P$, where P is a zero-ary predicates. However, in the canonical model for, say, FK+(Triv), we may well have $w \xrightarrow{\sigma} w'$ for some substitution σ and $w' \neq w$. It doesn't help to stipulate that σ must be bijective. For example, suppose w contains Fv_1, $\{\neg Fv_i : i \neq 1\}$, while w' contains Fv_2, $\{\neg Fv_i : i \neq 2\}$, and σ swaps v_1 and v_2. Then we may well have $w \xrightarrow{\sigma} w'$.

In this case, it would help to restrict the eligible substitutions to $\{id\}$. Obviously, however, we don't want to do this in general.

What we could do, perhaps, is use different types of canonical models for different systems. Canonical models for (B) systems would employ only bijective substitutions, while canonical models for (Triv) systems would employ only the identity substitution.

Unfortunately, this approach also doesn't work in general.

6.3 S4M

An interesting case study for the differences between propositional modal logics and their quantified extension is S4M (a.k.a. S4.1), which add the McKinsey axiom

(M) $\Box\Diamond A \supset \Diamond\Box A$

to S4 (= K+(T)+(4)).

Among transitive and reflexive Kripke frames, (M) defines the condition of *finality*: every world can see a world that can only see itself. But while propositional S4M is sound and complete with respect to the class of transitive, reflexive, and final frames (as shown in [Hughes and Cresswell 1996: pp.131–134]), quantified S4M is incomplete in Kripke semantics: it is not sound and complete with respect to any class of frames (as shown in [Hughes and Cresswell 1996: pp.265–270,283]).

The incompleteness proof (due to Fine) involves the following formula:

$$\Box \exists x Fx \supset \Diamond \exists x \Box Fx. \tag{$\star$}$$

This, it turns out, is not provable in QS4M (= QK+(T)+(4)+(M)), yet it is valid on every final Kripke frame. Since every frame for QS4M is final, it follows that QS4M is not complete with respect to any class of frames with respect to which it is sound.

Here, in outline, is the proof that $(\star)$ is valid on every final Kripke frame. Let $\mathfrak{F}$ be some such frame. Assume that

$$\mathfrak{F}, I, w, g \Vdash \Box \exists x Fx$$

for some I, w, g. We know that w can see a world w' that can only see itself. Thus

$$\mathfrak{F}, I, w', g \Vdash \exists x Fx.$$
$$\Rightarrow \text{exists } d \text{ s.t. } \mathfrak{F}, I, w', g^{x \mapsto d} \Vdash Fx$$
$$\Rightarrow \text{exists } d \text{ s.t. } \mathfrak{F}, I, w', g^{x \mapsto d} \Vdash \Box Fx$$
$$\Rightarrow \mathfrak{F}, I, w', g \Vdash \exists x \Box Fx$$
$$\Rightarrow \mathfrak{F}, I, w, g \Vdash \Diamond \exists x \Box Fx.$$

In counterpart semantics, finality turns into "local finality": for each $w, d_1, \ldots, d_n$ there is a sequence $w', d'_1, \ldots, d'_n$ such that $w, d_1, \ldots, d_n R^n w', d'_1, \ldots, d'_n$ and $w', d'_1, \ldots, d'_n$ is only R^n-related to itself. The above proof no longer goes through. Indeed, $(\star)$ is *not* valid on every locally final counterpart structure. For a counterexample, let $\mathfrak{S}$ be as follows.

$$W = \{w\},$$
$$R = \{(w, w)\},$$
$$U_w = D_w = \{a, b\},$$
$$K_{w,w} = \{\{(a, a), (b, b)\}, \{(a, b), (b, b)\}\}.$$

This structure is locally final, as the sequence $w, b, b, \ldots$ is an "end point" for any sequence $w, d_1, d_2, \ldots$ with $d_i \in U_w$. (There is a counterpart relation relative to which b is a counterpart of every individual, so we have $w, d_1, d_2, \ldots \triangleright w, b, b, \ldots$.

And since *b* is its only counterpart relative to both relations, $w, b, b, \ldots$ is $\rhd$-related only to itself.) The structure is also locally reflexive and transitive. But $(\star)$ is invalid on $\mathfrak{S}$. To see this, let $I(F, w) = \{a\}$ and let *g* be any assignment. Then $\mathfrak{S}, I, w, g \Vdash \Box\exists xFx$ because there are g' for which $w, g \rhd w, g'$, and $\mathfrak{S}, I, w, g' \Vdash \exists xFx$ for all such g'. But $\mathfrak{S}, I, w, g \not\Vdash \Diamond\exists x\Box Fx$. For $\Diamond\exists x\Box Fx$ is true at w, g iff some $\exists x\Box Fx$ is true at w, g' for some g' with $w, g \rhd w, g'$. And $\exists x\Box Fx$ is true at w, g' iff there is a *d* such that $\mathfrak{S}, I, w, g'^{x \mapsto d} \Vdash \Box Fx$, i.e., such that $\mathfrak{S}, I, w, g^* \Vdash Fx$ for all g^* with $w, g'^{x \mapsto d} \rhd w, g^*$. *d* can be either *a* or *b*. Both have *b* as a counterpart, relative to the second counterpart relation. So one eligible g^* always maps *x* to *b*. And then $\mathfrak{S}, I, w, g^* \not\Vdash Fx$.

So perhaps QS4M is complete in counterpart semantics. But we can't establish completeness by the method from the previous section. Here is why.

Since $(\star)$ is not in QS4M, some world *w* in the canonical model must contain its negation. So *w* contains $\Box\exists xFx$ and $\Box\neg\exists x\Box Fx$. Let *g* be an arbitrary assignment for *w*. If the structure of the canonical model is final, any $w, \underline{g}$ (for an initial segment $\underline{g}$ of *g*) can access a final point $w', \underline{g}'$, via some substitution σ. What does $w', \underline{g}'$ look like?

We must have $\exists xFx \in w'$, from $\Box\exists xFx \in w$. So there must be a witnessing $Fy \in w'$. But we can't have $\Box Fy \in w'$: otherwise $\exists x\Box Fx \in w'$ by existential generalisation, (recall that we're using a classical base in QS4M), which is incompatible with $\Box\neg\exists x\Box Fx \in w$ and $w, g \rhd w', g'$.

From Fy and $\neg\Box Fy$ in w', we can infer that there is an eligible substitution σ that leads from $w', [y]_{w'}$ to some $w'', [\sigma(y)]_{w''}$ such that $\neg F\sigma(y) \in w''$. Since $w', \underline{g}'$ is final, w'' must be w' itself. So there must be an eligible substitution σ such that $w' \xrightarrow{\sigma} w'$ and $\neg F\sigma(y) \in w'$.

Since $w', \underline{g}'$ can only see itself, it follows that $[y]_{w'}$ is not in $\underline{g}'$. (If it were then following the σ arrows would lead from $w', \underline{g}'$ to a different $w', \underline{g}''$.)

Since the structure is reflexive, there must be an arrow from $[y]_{w'}$ to itself. So there must be two arrows from $[y]_{w'}$, one to $[y]_{w'}$ and one to $[y']_{w'}$. $w', \underline{g}'$ is final, but $w', g'^{x \mapsto [y]_{w'}}$ is not. (In this respect, the structure looks like the simple example above.)

Now here's the problem. Why can't w' see another world w'' via the same substitution σ, which would make $w', \underline{g}'$ non-final?

I haven't yet said much about what $w', \underline{g}'$ looks like. Recall that we only need to

show that for w, g can access *some* final w', g'. We can construct these however we want. Presumably we'll use the fact that QS4M can prove $\Diamond(A \to \Box A)$, for any finite conjunction of $A \to \Box A$ formulas. It follows that

$$\Gamma = \{\tau(B) : \Box B \in w\} \cup \{\tau(A \supset \Box A) : A \text{ a sentence}\}$$

is QS4M-consistent, as long as τ is injective (so that it has an inverse). Now we'd try to construct w' out of Γ and g' out of τ and g. w' would then contain all $A \supset \Box A$ with $\text{Var}(A) \in \text{Ran}(\tau)$. Since w' contains Fx for all $x \in g'$, it would presumably contain $\Box Fx$ for all these x. (We can't require that w' contain $A \supset \Box A$ for *all* A: we know that it contains Fy but not $\Box Fy$.) But this isn't enough to ensure that w' can't see another world w'' via σ. It only tells us that this other world matches w' with respect to sentences whose variables are in g'. For example, w'' must also have $\neg Fx$. But why can't we swap two of the other individuals – say, the witness $[y]_{w'}$ with $Fy \in w'$ and a non-witness $[y']_{w'}$ with $\neg Fy' \in w'$, so that in w'' we have Fy' and $\neg Fy$?

7 Non-functional logics

7.1 Multiple counterparts and contingent identity

The logics we have studied so far have been well-behaved, displaying none of the deviant features of Lewis's logic. This docility had two sources. One was our strong reading of the box. The other was our assumption of functionality: we have assumed that every individual has at most one counterpart at every accessible world, relative to any counterpart relation. We are now going to relax this assumption.

Why should we do this? One possible motivation is that we may want to allow not only for contingent distinctness, but also for contingent identity. Functional counterpart semantics validates (NI) but not (ND).

(NI) $x = y \supset \Box(x = x \supset x = y)$.

(ND) $x \neq y \supset \Box x \neq y$.

One might want to treat these on a par and allow both to fail.

Remember, however, that we treat singular terms as purely referential: the compositional semantic value of a name or variable is exhausted by its referent. Putative examples of contingent identity, as in [Gibbard 1975], often involve terms ('Lumpl', 'Goliath') that are associated with different ways of tracking an individual across worlds. We could add such an association to our semantics, but we have not done so.

That said, there are possible examples of (metaphysically) contingent identity and distinctness that do not involve different ways of tracking individuals. Consider a group of small islands, formed as the result of volcanic activity. We might want to say that two of these islands could have been a single island, if a little more lava had flowed in between them. (Compare [Karmo 1983].) Conversely, we might want to say that a single island could have been two islands, if a little less lava had flowed into its central valley. (Compare Lewis's [1973: 40f.] intuition that he "might have

79

been twins". [Schwarz 2014] considers analogous temporal cases. Epistemic cases are easy to construct.)

Unlike (ND), (NI) is provable by standard formulations of Leibniz's Law. This is sometimes said to show that (NI) is a logical truth. But you can't just look at the shape of a formula to see whether it is a proper instance of Leibniz's Law. If we give the box the semantics of $\forall y$, then (NI) is equivalent to the following, evidently invalid, statement:

$$x = y \supset (\forall y\, x = x \supset \forall y\, x = y).$$

This is not a proper instance of Leibniz's Law because the variable y gets captured when the second occurrence of x in $\forall y\, x = x$ is replaced by y. As Lewis [1983] observed, in counterpart semantics, modal operators effectively function as unselective binders of all variables in their scope. Informally, if x and y denote the same individual, then $\Box x = x$ says that all counterparts of this individual are self-identical, but $\Box x = y$ says that all counterparts of the individual are identical *to one another*.

An adequate logic for non-functional counterpart structures must restrict the application of substitution principles. But how? Since terms as purely referential, modal contexts are not *generally* resistant to substitution. The following principle, for example, is valid:

$$x = y \supset (\Diamond Fx \supset \Diamond Fy).$$

In the next section, I am going to explain the restrictions we'll need.

7.2 Substitution revisited

In section 3.2, I reviewed three strategies to prevent unwanted capturing of variables in substitution. Let's go over them again.

One strategy is to restrict substitution principles so that the substituted term must be an individual constant. This wouldn't help in the present context because modal operators are unselective binders of *singular terms*, not just of variables: (NI) is invalid in non-functional structures even if x and y are names.

A second strategy is to redefine the substitution operation so that bound occurrences of variables get renamed before substitution. We are going to see that this route is also blocked: there is no way to redefine the substitution operation $[y/x]$ that satisfies the substitution lemma in non-functional structures.

This leaves us with the third option: we have to explicitly restrict substitution principles. Leibniz' Law, for example, will become the schema

(LL*) $x = y \supset (A \supset [y/x]A)$, provided y is free for x in A.

In classical first-order logic, a variable y is *free (to be substituted) for x in A* iff no free occurrence of x in A lies in the scope of a quantifier that binds y. The same restriction works for functional counterpart structures. For non-functional structures, we need a stronger concept of freedom for substitution. The following (somewhat non-obvious) definition turns out to work.

Definition 7.1 (Modal separation and freedom for substitution).
Two variables x and y are *modally separated in* a formula A if no free occurrences of x and y in A lie in the scope of the same modal operator.

y is *really free (to be substituted) for x in A* if either (i) x and y are the same variable, or (ii) x and y are modally separated in A, or (iii) A has the form $\Box B$ and y is really free for x in B.

Clause (iii) makes the definition recursive.

I say '*really* free', rather than 'free', as a reminder that this is not the usual notion of freedom for substitution that you may be familiar with from first-order logic.

To illustrate definition 7.1, x and y are modally separated in $\Diamond Fx$ and $\Box Fx \supset \Diamond Fy$ and $\forall x \Box Gxy$; y is really free for x in $\Box x = y$ and $\Box\Box\neg Gxy$ and $\Box\Diamond\neg\exists x Gxy$, but not in $\Box\Diamond\neg Gxy$. Accordingly,

$$x = y \supset (\Diamond Fx \supset \Diamond Fy) \text{ and}$$
$$x = y \supset (\Box x = y \supset \Box y = y) \text{ and}$$
$$x = y \supset (\Diamond x = y \supset \Box y = y) \text{ and}$$
$$x = y \supset (\Box\Box\neg Gxy \supset \Box\Box\neg Gyy) \text{ and}$$
$$x = y \supset (\Box\Box\neg\exists x Gxy \supset \Box\Box\neg\exists z Gyz)$$

are valid, but

$$x = y \supset (\Box\Diamond\neg Gxy \supset \Box\Diamond\neg Gyy)$$

is invalid.

Definition 7.1 does not incorporate the traditional restriction, that no free occurrence of x in A lies in the scope of a quantifier that binds y. That's because I assume that $[y/x]$ still renames overtly bound variables, as per definition 3.1. We could, of course, have used a naive definition of substitution in substitution principles like (LL*) and added the traditional restriction to the conditions under which y is (really) free for x in A.

I'm now going to prove two key facts about modal separation and real freedom for substitution. First, modal separation ensures that the substitution lemma holds:

Lemma 7.1 (Substitution lemma with modal separation).
For any counterpart model $\mathfrak{M}$, world w in $\mathfrak{M}$, assignment g on U_w, sentence A, and variables x and y,

$$\mathfrak{M}, w, g^{[y/x]} \Vdash A \text{ iff } \mathfrak{M}, w, g \Vdash [y/x]A,$$

provided x and y are modally separated in A.

Recall that $g^{[y/x]}$ (a.k.a. $g \circ [y/x]$) is the x-variant of g that assigns $g(y)$ to x.

Proof. The proof is by induction on A. All cases is except the one for $A = \Box B$ are covered in the proof of lemma 3.1, with $\sigma = [y/x]$. In each case, we assume that x and y are modally separated in A. It follows that x and y are also modally separated in the immediate subformulas B and C or A, that the induction hypotheses in the proof of lemma 3.1 apply. The only exception is the case for $A = \Box B$, which we now consider. We can assume that x and y are different variables: the target claim trivially holds otherwise.

By definition 2.9, $\mathfrak{M}, w, g^{[y/x]} \Vdash \Box B$ iff $\mathfrak{M}, w', g^{[y/x]\prime} \Vdash B$ for all $w', g^{[y/x]\prime}$ with $w, g^{[y/x]} R w', g^{[y/x]\prime}$. We also have $\mathfrak{M}, w, g \Vdash [y/x]\Box B$ iff $\mathfrak{M}, w, g \Vdash \Box[y/x]B$ by definition 3.1, iff $\mathfrak{M}, w', g' \Vdash [y/x]B$ for all w', g' with $w, g \triangleright w', g'$, by definition 2.9, iff $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$ for all such w', g' by induction hypothesis. So we have to show that

$$\mathfrak{M}, w', g^{[y/x]\prime} \Vdash B \text{ for all } w', g^{[y/x]\prime} \text{ with } w, g^{[y/x]} \triangleright w', g^{[y/x]\prime} \tag{1}$$

iff

$$\mathfrak{M}, w', g'^{[y/x]} \Vdash B \text{ for all } w', g' \text{ with } w, g \triangleright w', g'. \qquad (2)$$

(1) entails (2). For assume w', g' are such that $w, g \triangleright w', g'$. By definition 2.8, this means that there is a $C \in K_{w,w'}$ such that $g(x)Cg'(x)$ for all variables x. So we have, for all variables x, $g([y/x]x)Cg'([y/x]x)$; in other words: $g^{[y/x]}(x)Cg'^{[y/x]}(x)$. So $w, g^{[y/x]} \triangleright w', g'^{[y/x]}$. Using $g'^{[y/x]}$ as $g^{[y/x]'}$ it follows by (1) that $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$.

Without further assumptions, however, the converse direction may fail. $g'^{[y/x]}$ and $g^{[y/x]'}$ assign to x and y some counterpart of $g(y)$ (if there is any). But while $g'^{[y/x]}$ assigns *the same* counterpart to x and y, $g^{[y/x]'}$ may choose different counterparts for x and y, relative to the same counterpart relation.

But we know that either x or y does not occur freely in B. To show that (2) entails (1), assume that $w', g^{[y/x]'}$ are such that $w, g^{[y/x]} \triangleright w', g^{[y/x]'}$. By definition 2.8, this means that there is a $C \in K_{w,w'}$ such that $g^{[y/x]}(z)Cg^{[y/x]'}(z)$, for all variables z.

Now assume that x is not free in B. Define g' so that $g'(z) = g^{[y/x]'}(z)$ for every variable z other than x, and $g'(x)$ is an arbitrary d with $g(x)Cd$, or undefined if there is no such d. Then $g(z)Cg'(z)$ for all variables z, and so $w, g \triangleright w', g'$. By (2), we have $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$. But $g'^{[y/x]}$ and $g^{[y/x]'}$ differ only in the value of x, which is not free in B. By lemma 2.2, it follows that $\mathfrak{M}, w', g^{[y/x]'} \Vdash B$.

Next, assume that y is not free in B. In this case, define g' so that $g'(x)$ is an arbitrary C-counterpart of $g(x)$, as before, $g'(y) = g^{[y/x]'}(x)$, and $g'(z) = g^{[y/x]'}(z)$ for every variable z other than x and y. Again, we have $g(z)Cg'(z)$ for all variables z, and thus $w, g \triangleright w', g'$. (Note that $g^{[y/x]'}(x)$ is a C-counterpart of $g(y)$.) By (2), $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$. But $g'^{[y/x]}$ and $g^{[y/x]'}$ differ only in the value of y, which is not free in B. By lemma 2.2, it follows that $\mathfrak{M}, w', g^{[y/x]'} \Vdash B$.

If we weaken the restriction of modal separation to real freedom of y for x, the substitution lemma holds only in one direction:

Lemma 7.2 (Semi-substitution lemma).
For any counterpart model $\mathfrak{M}$, world w in $\mathfrak{M}$, assignment g on U_w, and sen-

tence A,

$$\text{if } \mathfrak{M}, w, g^{[y/x]} \Vdash A \text{ then } \mathfrak{M}, w, g \Vdash [y/x]A,$$

provided y is really free for x in A.

Proof. The target claim obviously holds if x and y are the same variable. Assume they are not. We proceed by induction on A, but we only need two cases.

First, A is not of the form $\Box B$. In that case, y is really free for x in A iff x and y are modally separated in A, by definition 7.1. The target claim then follows from lemma 7.1.

Assume then that A is $\Box B$. We assume that $\mathfrak{M}, w, g^{[y/x]} \Vdash \Box B$, and try to derive $\mathfrak{M}, w, g \Vdash [y/x]\Box B$, provided y is really free for x in $\Box B$.

By definition 2.9, $\mathfrak{M}, w, g^{[y/x]} \Vdash \Box B$ implies that $\mathfrak{M}, w', g^{[y/x]\prime} \Vdash B$ for all $w', g^{[y/x]\prime}$ with $w, g^{[y/x]} R w', g^{[y/x]\prime}$.

Let w', g' be such that $w, g \triangleright w', g'$. By definition 2.8, this means that there is a $C \in K_{w,w'}$ such that $g(z) C g'(z)$ for all variables z. So we also have $g([y/x]z) C g'([y/x]z)$; in other words: $g^{[y/x]}(z) C g'^{[y/x]}(z)$, for all z. So $w, g^{[y/x]} \triangleright w', g'^{[y/x]}$. Since $\mathfrak{M}, w', g^{[y/x]\prime} \Vdash B$ for all $w', g^{[y/x]\prime}$ with $w, g^{[y/x]} R w', g^{[y/x]\prime}$, we can infer that $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$.

Assume that y is really free for x in $\Box B$. By definition 7.1, it follow that y is really free for x in B. So by induction hypothesis, if $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$ then $\mathfrak{M}, w', g' \Vdash [y/x]B$. We have shown that $\mathfrak{M}, w', g'^{[y/x]} \Vdash B$. So $\mathfrak{M}, w', g' \Vdash [y/x]B$.

Since w', g' were arbitrary points with $w, g \triangleright w', g'$, we have $\mathfrak{M}, w, g \Vdash \Box[y/x]B$ by definition 2.9, and thus $\mathfrak{M}, w, g \Vdash [y/x]\Box B$ by definition 3.1.

The converse does not hold. For example, $\mathfrak{M}, w, g \Vdash [y/x]\Box x = y$ does not imply $\mathfrak{M}, w, g^{[y/x]} \Vdash \Box x = y$. The operation $[y/x]$, as defined in definition 3.1, does not always satisfy the substitution lemma, not even when y is really free for x. The following proposition shows that this problem is unavoidable.

Proposition 7.3 (Undefinability of substitution).
There is no operation Φ on formulas A in the standard language of modal

> *predicate logic such that $\mathfrak{M}, w, g \Vdash \Phi(A)$ iff $\mathfrak{M}, w, g^{[y/x]} \Vdash A$ for all models $\mathfrak{M}$, worlds w in $\mathfrak{M}$, assignments g on U_w, and variables x, y.*

Proof. We are going to show that the standard language of modal predicate logic cannot express that an individual has multiple counterparts at some accessible world (relative to the same counterpart relation).

Let $\mathfrak{M}_1$ be a counterpart model with $W = \{w\}$, $R = \{\langle w, w \rangle\}$, $U_w = \{x, y, y^*\}$, $D_w = \{x\}$, $K_{w,w} = \{\{\langle d, d \rangle : d \in U_w\}\}$, and $I_w(P) = \emptyset$ for all non-logical predicates P. Let $g(y) = y$ and $g(z) = x$ for every variable $z \neq y$. Let $\mathfrak{M}_2$ be like $\mathfrak{M}_1$ except that y^* is also a counterpart of y. That is, $K_{w,w'} = \{\{\langle x, x \rangle, \langle y, y \rangle, \langle y^*, y^* \rangle, \langle y, y^* \rangle\}\}$.

We have $\mathfrak{M}_2, w, g^{[y/x]} \Vdash \Diamond y \neq x$, but $\mathfrak{M}_1, w, g^{[y/x]} \not\Vdash \Diamond y \neq x$. However, every $\mathfrak{L}$-sentence has the same truth-value at w under g in both models.

We prove this by showing that for every $\mathfrak{L}$-sentence A, the following three statements are equivalent: (1) $\mathfrak{M}_1, w, g \Vdash A$, (2) $\mathfrak{M}_2, w, g \Vdash A$, (3) $\mathfrak{M}_2, w, g^* \Vdash A$, where g^* is the y-variant of g on U_w with $g^*(y) = g(y^*)$.

1. *A is $Px_1 \ldots x_n$.* It is clear that $\mathfrak{M}_1, w, g \Vdash Px_1 \ldots x_n$ iff $\mathfrak{M}_2, w, g \Vdash Px_1 \ldots x_n$ because counterpart relations do not figure in the evaluation of atomic formulas. Moreover, for non-logical P, $\mathfrak{M}_2, w, g \not\Vdash Px_1 \ldots x_n$ and $\mathfrak{M}_2, w, g^* \not\Vdash Px_1 \ldots x_n$, because $I(P) = \emptyset$. For the identity predicate, observe that $\mathfrak{M}_2, w, g \not\Vdash u = v$ iff exactly one of u, v is y, since $g(z) = x$ for all terms $z \neq y$. For the same reason, $\mathfrak{M}_2, w, g^* \not\Vdash u = v$ iff exactly one of u, v is y. So $\mathfrak{M}_2, w, g \Vdash u = v$ iff $\mathfrak{M}_2, w, g^* \Vdash u = v$.

2. *A is $\neg B$.* $\mathfrak{M}_1, w, g \Vdash \neg B$ iff $\mathfrak{M}_1, w, g \not\Vdash B$ by definition 2.9, iff $\mathfrak{M}_2, w, g \not\Vdash B$ by induction hypothesis, iff $\mathfrak{M}_2, w, g \Vdash \neg B$ by definition 2.9. Moreover, $\mathfrak{M}_2, w, g \not\Vdash B$ iff $\mathfrak{M}_2, w, g^* \not\Vdash B$ by induction hypothesis, iff $\mathfrak{M}_2, w, g^* \Vdash \neg B$ by definition 2.9.

3. *A is $B \supset C$.* Analogous.

4. *A is $\forall z B$.* Let v be a variable not in $\mathrm{Var}(B) \cup \{y\}$. By lemma 7.1, $\mathfrak{M}_1, w, g \Vdash \forall z B$ iff $\mathfrak{M}_1, w, g \Vdash \forall v [v/z] B$. By definition 2.9, $\mathfrak{M}_1, w, g \Vdash \forall v [v/z] B$

iff $\mathfrak{M}_1, w, g^{v \mapsto d} \Vdash [v/z]B$ for all $d \in D_w$. As $D_w = \{x\}$ and $g(v) = x$, the only such $g^{v \mapsto d}$ is g itself. So $\mathfrak{M}_1, w, g \Vdash \forall z B$ iff $\mathfrak{M}_1, w, g \Vdash [v/z]B$. By the same reasoning, $\mathfrak{M}_2, w, g \Vdash \forall z B$ iff $\mathfrak{M}_2, w, g \Vdash [v/z]B$. But by induction hypothesis, $\mathfrak{M}_1, w, g \Vdash [v/z]B$ iff $\mathfrak{M}_2, w, g \Vdash [v/z]B$. So $\mathfrak{M}_1, w, g \Vdash \forall z B$ iff $\mathfrak{M}_2, w, g \Vdash \forall z B$. Moreover, by induction hypothesis, $\mathfrak{M}_2, w, g \Vdash [v/z]B$ iff $\mathfrak{M}_2, w, g^* \Vdash [v/z]B$, iff $\mathfrak{M}_2, w, g^* \Vdash \forall v[v/z]B$ because g^* is the only $g^{v \mapsto d}$ with $d \in D_w$, iff $\mathfrak{M}_2, w, g^* \Vdash \forall z B$ by lemma 7.1.

5. A is $\Box B$. In both $\mathfrak{M}_1$ and $\mathfrak{M}_2$, the only world accessible from w is w itself. In $\mathfrak{M}_1$, g at w is also the only image of g at w. So by definition 2.9, $\mathfrak{M}_1, w, g \Vdash \Box B$ iff $\mathfrak{M}_1, w, g \Vdash B$.

   In $\mathfrak{M}_2$, there are two w-images of g at w: g and g^*. So $\mathfrak{M}_2, w, g \Vdash \Box B$ iff both $\mathfrak{M}_2, w, g \Vdash B$ and $\mathfrak{M}_2, w, g^* \Vdash B$. By induction hypothesis, $\mathfrak{M}_1, w, g \Vdash B$ iff both $\mathfrak{M}_2, w, g \Vdash B$ and $\mathfrak{M}_2, w, g^* \Vdash B$. So $\mathfrak{M}_1, w, g \Vdash \Box B$ iff $\mathfrak{M}_2, w, g \Vdash \Box B$. Moreover, in $\mathfrak{M}_2$, g^* is the only w-image of g^* at w. So $\mathfrak{M}_2, w, g^* \Vdash \Box B$ iff $\mathfrak{M}_2, w, g^* \Vdash B$. By induction hypothesis, $\mathfrak{M}_2, w, g^* \Vdash B$ iff $\mathfrak{M}_2, w, g \Vdash B$. So $\mathfrak{M}_2, w, g^* \Vdash \Box B$ iff both $\mathfrak{M}_2, w, g^* \Vdash B$ and $\mathfrak{M}_2, w, g \Vdash B$, which as we just saw holds iff $\mathfrak{M}_2, w, g \Vdash \Box B$.

7.3 The new base logics

We know that among functional counterpart structures, FK is the logic of total structures, NK is the logic of single-domain structures, and QK is the logic of structures that are both total and single-domain. This picture remains essentially intact for nonfunctional structures, except that all substitution principles of (free or classical) first-order logic must be restricted by the condition that the substituted term be "really free" for the term it replaces. Here are the revised principles.

(UI*) $\forall x A \supset [y/x]A$, provided y is really free for x in A,

(FUI*) $\forall x A \supset (Ey \supset [y/x]A)$, provided y is really free for x in A,

(LL*) $x = y \supset A \supset [y/x]A$, provided y is really free for x in A,

(Sub*) if $\vdash_L A$, then $\vdash_L [y/x]A$, provided y is really free for x in A.

Let FK* be the smallest system that contains all $\mathcal{L}$-instances of (Taut), (UD), (VQ), (FUI*), ($\forall$Ex), (=R), (LL*), (K), and that is closed under (MP), (UG), (Nec) and (Sub*).

Let QK* be the smallest system that contains all $\mathcal{L}$-instances of (Taut), (UD), (UI*), ($\forall$Ex), (=R), (LL*), (K), and that is closed under (MP), (UG), (Nec) and (Sub*).

Let NK* be the smallest system that contains all $\mathcal{L}$-instances of (Taut), (UD), (VQ), (FUI*), (Neg), (LL*), ($\forall$=R), (K), (NA), (TE), and that is closed under (MP), (UG), (Nec) and (Sub*).

> **Lemma 7.4 (Soundness of the FK* axioms).**
> *Every instance of* (Taut), (VQ), (UD), (FUI*), ($\forall$E!), (=R), (LL*), *and* (K) *is valid on every total counterpart structure.*

Proof. Let $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ be any total counterpart model, w a world in W, and g a (total) assignment on U_w. We show that $\mathfrak{M}, w, g \Vdash A$ for every instance A of every axiom.

The proof for (Taut), (VQ), (UD), ($\forall$E!), (=R), (K) is exactly as in the proof of lemma 3.3. We only have to adjust the proofs for (FUI*) and (LL*), where we'll invoke lemma 7.2 instead of lemma 3.1.

1. (FUI*). Assume $\mathfrak{M}, w, g \Vdash \forall x A$ and $\mathfrak{M}, w, g \Vdash E!t$. By lemma 3.2, the latter means that $g(t) \in D_w$. By definition 2.9, the former means that $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for every $d \in D_w$. So in particular, $\mathfrak{M}, w, g^{x \mapsto g(t)} \Vdash A$. Since $g^{x \mapsto g(t)} = g^{[t/x]}$, it follows by lemma 7.2 that $\mathfrak{M}, w, g \Vdash [t/x]A$.

2. (LL*). Assume $\mathfrak{M}, w, g \Vdash s = t$ and $\mathfrak{M}, w, g \Vdash A$. By definitions 2.9 and 2.4, the former implies that $g(s) = g(t)$. So $g^{[s/t]} = g$. Since $\mathfrak{M}, w, g \Vdash A$, we have $\mathfrak{M}, w, g^{[s/t]} \Vdash A$. It follows by lemma 7.2 that $\mathfrak{M}, w, g \Vdash [s/t]A$.

Lemma 7.5 (Soundness of the FK* rules).
If all elements of a set of formulas Γ are valid on a counterpart structure $\mathfrak{S}$, and Γ is extended by (MP)*,* (UG) *or* (Nec)*, or* (Sub*)*, then the new sentences are still valid on $\mathfrak{S}$.*

Proof. The case of (MP), (UG), and (Nec) is covered in the proof of lemma 3.4. We only need to adjust the proof for (Sub*).

Assume for contraposition that $\mathfrak{M}, w, g \nVdash [y/x]A$ for some $\mathfrak{M}, w, g$. By lemma 7.2, it follows that $\mathfrak{M}, w, g \circ [y/x] \nVdash A$. Since $g \circ [y/x]$ is an assignment on U_w, this means that A is not valid on the structure of $\mathfrak{M}$.

Theorem 7.6 (Soundness of FK*).
All members of FK* are valid on all total counterpart structures.

Proof. Immediate from lemmas 7.4 and 7.5.

Lemma 7.7 (Soundness of (UI*)).
All instances of (UI*) *are valid on every total single-domain counterpart structure.*

Proof. Assume $\mathfrak{M}, w, g \Vdash \forall x A$. By definition 2.9, $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for every $d \in D_w$. Since g is total, $g(t) \in U_w$. So $g(t) \in D_w$. And so $\mathfrak{M}, w, g^{x \mapsto g(t)} \Vdash A$. Since $g^{x \mapsto g(t)} = g^{[t/x]}$, it follows by lemma 7.2 that $\mathfrak{M}, w, g \Vdash [t/x]A$.

Theorem 7.8 (Soundness of QK*).
All members of QK* are valid on all classical counterpart structures.

Proof. Immediate from lemmas 7.4, 7.5 and 7.7.

For negative logics, we use the concept of n-validity from section 3.3, as we allow

for terms that go genuinely empty.

Lemma 7.9 (Soundness of the NK* **axioms).**
Every instance of (Taut)*,* (VQ)*,* (UD)*,* (FUI*)*,* (Neg)*,* ($\forall = $R)*,* (LL*)*,* (NA)*,* (TE)*, and* (K) *is n-valid on every single-domain counterpart structure.*

Proof. The proof for all the axioms proceeds just as in the proofs of lemma 3.3, 3.8, and 7.4.

Theorem 7.10 (Soundness of NK*).**
Every member of NK* *is n-valid on every single-domain counterpart structure.*

Proof. Immediate from lemmas 7.9 and 7.5.

The completeness of these logics will be established in chapter 9.

7.4 Some consequences

Let's prove a few properties derivable from the above axiomatisations. We redefine the concept of positive and negative logics: from now on, a positive or negative logic only requires the weakened substitution principles.

Definition 7.2 (Positive logics).
A set of $\mathfrak{L}$-sentences is a *positive logic* if it includes FK* and is closed under (MP), (UG), (Nec) and (Sub*).

Definition 7.3 (Negative logics).
A set of $\mathfrak{L}$-sentences is a *negative logic* if it includes NK* and is closed under (MP), (UG), (Nec) and (Sub*).

Now let *L* be an arbitrary positive or negative logic. Lemmas 3.10 and 3.11 can be established just as in section 3.4. The same is true for lemmas 3.12, 3.13, and 3.15, except that we need to use the starred substitution principles. Here are the adjusted proofs of lemmas 3.12 and 3.13, for future reference.

Lemma 7.11 (Existence and self-identity).
If L is negative, *then for any $\mathfrak{L}$-variable x,*

  (EI) $\vdash_L E!x \leftrightarrow x = x;$

Proof. By (FUI*), $\vdash_L \forall x(x = x) \supset (E!x \supset x = x)$. By ($\forall = R$), $\vdash_L \forall x(x = x)$. So $\vdash_L E!x \supset x = x$. Conversely, $x = x \supset E!x$ by (Neg). $\square$

Lemma 7.12 (Symmetry and transitivity of identity).
For any $\mathfrak{L}$-variables x, y, z,

  ($= S$) $\vdash_L x = y \supset y = x;$
  ($= T$) $\vdash_L x = y \supset y = z \supset x = z.$

Proof. For (= S), let *v* be some variable $\notin \{x, y\}$. Then

  1. $\vdash_L v = y \supset (v = x \supset y = x)$. (LL*)
  2. $\vdash_L x = y \supset (x = x \supset y = x)$. (1, (Sub*))
  3. $\vdash_L x = y \supset x = x$. (($= R$), or (Neg) and ($\forall = R$))
  4. $\vdash_L x = y \supset y = x$. (2, 3)

For (= T),

  1. $\vdash_L x = y \supset y = x$. (= S)
  2. $\vdash_L y = x \supset (y = z \supset x = z)$. (LL*)
  3. $\vdash_L x = y \supset (y = z \supset x = z)$. (1, 2)

90

> **Lemma 7.13 (Closure under injective substitutions).**
> *For any $\mathfrak{L}$-formula A and injective substitution τ on $\mathfrak{L}$,*
>
> (Sub^t) $\vdash_L A$ *iff* $\vdash_L A^\tau$.

Proof. Assume $\vdash_L A$. Let $x_1, \dots, x_n$ be the variables in A. If $n = 0$, then $A = A^\tau$ and the result is trivial. If $n = 1$, then A^τ is $[x_1^\tau/x_1]A$, and x_1^τ is either x_1 itself or does not occur in A. In the first case, $[x_1^\tau/x_1]A = A$ and the result is again trivial. In the second case, x_1^τ is really free for x_1 in A, and thus $\vdash_L [x_1^\tau/x_1]A$ by (Sub^*).

Assume then that $n > 1$. Note first that $A^\tau = [x_n^\tau/v_n] \dots [x_2^\tau/v_2][x_1^\tau/x_1]$ $[v_2/x_2] \dots [v_n/x_n]A$, where $v_2, \dots, v_n$ are distinct variables not in A or A^τ. This is easily shown by induction on the subformulas B of A (ordered by complexity). To keep lines short, let Σ abbreviate $[x_n^\tau/v_n] \dots [x_2^\tau/v_2][x_1^\tau/x_1][v_2/x_2] \dots [v_n/x_n]$.

1. If B is $Px_j \dots x_k$, then $x_j, \dots, x_k$ are variables from $x_1, \dots, x_n$, and $\Sigma B = Px_j^\tau \dots x_k^\tau = B^\tau$, by definition 3.1.

2. If B is $\neg C$, then by induction hypothesis, $\Sigma C = C^\tau$, and hence $\neg \Sigma C = \neg C^\tau$. But $\Sigma \neg C$ is $\neg \Sigma C$ by definition 3.1, and $(\neg C)^\tau$ is $\neg C^\tau$ by definition 3.1.

3. The case for $C \supset D$ is analogous.

4. If B is $\forall z C$, then by induction hypothesis, $\Sigma C = C^\tau$. Since τ is injective, $\Sigma \forall z C$ is $\forall \Sigma z \Sigma C$ by definition 3.1, and $(\forall z C)^\tau$ is $\forall z^\tau C^\tau$ by definition 3.1. Moreover, since z is one of $x_1, \dots, x_n$, $\Sigma z = z^\tau$.

5. If B is $\Box C$, then by induction hypothesis, ΣC is C^τ, and hence $\Box \Sigma C$ is $\Box C^\tau$. But $\Sigma \Box C$ is $\Box \Sigma C$ by definition 3.1, and $(\Box C)^\tau$ is $\Box C^\tau$ by definition 3.1.

Now we show that L contains all "segments" of $[x_n^\tau/v_n] \dots [x_2^\tau/v_2][x_1^\tau/x_1]$ $[v_2/x_2] \dots [v_n/x_n]A$, beginning with the rightmost substitution, $[v_n/x_n]A$. Since

v_n is really free for x_n in A, by (Sub*), $\vdash_L [v_n/x_n]A$. Likewise, for each $1 < i < n$, v_i is really free for x_i in $[v_{i+1}/x_{i+1}]\dots[v_n/x_n]A$. So $\vdash_L [v_2/x_2]\dots[v_n/x_n]A$.

With respect to $[x_1^\tau/x_1]$, we distinguish three cases. First, if $x_1 = x_1^\tau$, then $\vdash_L [x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$, because $[x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$ is $[v_2/x_2]\dots[v_n/x_n]A$. Second, if $x_1 \neq x_1^\tau$ and $x_1^\tau \notin \mathrm{Var}(A)$, then $x_1^\tau \notin \mathrm{Var}([v_2/x_2]\dots[v_n/x_n]A)$, since the $v_1,\dots,v_n$ are not in $\mathrm{Var}(A)$ or $\mathrm{Var}(A^\tau)$ So x_1^τ is really free for x_1 in $[v_2/x_2]\dots[v_n/x_n]A$, and by (Sub*), $\vdash_L [x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$. Third, if $x_1 \neq x_1^\tau$ and $x_1^\tau \in \mathrm{Var}(A)$, then x_1^τ must be one of $x_2,\dots,x_n$. Then again $x_1^\tau \notin \mathrm{Var}([v_2/x_2]\dots[v_n/x_n]A)$, and so $\vdash_L [x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$ by (Sub*).

Next, x_2^τ is really free for v_2 in $[x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$, because τ is injective and hence $x_2^\tau \neq x_1^\tau$, so x_2^τ does not occur in $[x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$. Hence $\vdash_L [x_2^\tau/v_2][x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$. By the same reasoning, for each $2 < i \leq n$, x_i^τ is really free for v_i in $[x_{i-1}^\tau/v_{i-1}]\dots[x_2^\tau/v_2][x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$. So $\vdash_L [x_n^\tau/v_n]\dots[x_2^\tau/v_2][x_1^\tau/x_1][v_2/x_2]\dots[v_n/x_n]A$, i.e. $\vdash_L A^\tau$.

This proves the left-to-right direction of (Sub^t). The other direction immediately follows. Let $x_1^\tau,\dots,x_n^\tau$ be the variables in A^τ, and let σ be an arbitrary transformation that maps each x_i^τ back to x_i (i.e., to $(x_i^\tau)^{\tau^{-1}}$). By the left-to-right direction of (Sub^t), $\vdash_L A^\tau$ entails $\vdash_L (A^\tau)^\sigma$, and $(A^\tau)^\sigma$ is simply A.

Lemma 7.14 (Leibniz' Law with partial substitution).
Let A be a formula of $\mathfrak{L}$, and x, y variables of $\mathfrak{L}$. Let $[y//x]A$ be A with one or more *occurrences of x replaced by y.*

(LL$_\mathrm{p}^*$) $\vdash_L x = y \supset A \supset [y//x]A$, *provided the following conditions are all satisfied.*
- (i) $[y//x]A$ *does not replace any occurrence of x in the scope of a quantifier binding x or y.*
- (ii) *Either y is really free for x in A, or $[y//x]A$ does not replace any occurrence of x in the scope of a modal operator in A that also contains y.*
- (iii) *In the scope of any modal operator in A, $[y//x]A$ either replaces all or no occurrences of x by y.*

Proof. Let $v \neq y$ be a variable not in $\mathrm{Var}(A)$, and let $[v//x]A$ be like $[y//x]A$ except that all new occurrences of y are replaced by v: if $[y//x]A$ satisfies (i)–(iii), then so does $[y//x]A$ with all new occurrences of y replaced by v. Moreover, in the resulting formula $[v//x]A$ all occurrences of v are free and free for y, by clause (i); so $[y/v][v//x]A = [y//x]A$ by definition 3.1. By (LL*),

$$\vdash_L v = y \supset [v//x]A \supset [y/v][v//x]A, \tag{1}$$

provided that y is really free for v in $[v//x]A$, i.e. provided that either y is really free for x in A, or $[v//x]A$ (and thus $[y//x]A$) does not replace any occurrence of x in the scope of a modal operator in A that also contains y. This is guaranteed by condition (ii). Since $[y/v][v//x]A$ is $[y//x]A$, (1) can be shortened to

$$\vdash_L v = y \supset [v//x]A \supset [y//x]A. \tag{2}$$

By (Sub*), it follows that

$$\vdash_L [x/v](v = y \supset [v//x]A \supset [y//x]A), \tag{3}$$

provided that x is really free for v in $v = y \supset [v//x]A \supset [y//x]A$. Since this isn't a formula of the form $\Box B$, x is really free for v here iff no free occurrences of x and v lie in the scope of the same modal operator in $[v//x]A$. So whenever $[v//x]A$ (and thus $[y//x]A$) replaces some occurrences of x in the scope of a modal operator in A, then it must replace all occurrences of x in the scope of that operator. This is guaranteed by condition (iii). By definition 3.1, (3) can be simplified to

$$\vdash_L x = y \supset A \supset [y//x]A. \tag{4}$$

I will never actually use $(\mathrm{LL}_\mathrm{p}^*)$. I mention it only because Leibniz' Law is often stated for partial substitutions, and you may have wondered what that would look like in our systems. Now you know. We could indeed have used $(\mathrm{LL}_\mathrm{p}^*)$ as basic axiom instead of (LL*); (LL*) would then be derivable, because (a) every formula A has an alphabetic variant A' such that $[y/x]A$ is an instance of $[y//x]A'$ that satisfies (i)–(iii) iff y is really free for x in A, and (b) (LL*) is not used in the proof of lemma

3.15. I have chosen (LL*) as basic due to its much greater simplicity.

(The axiomatization in [Kutz 2000: 43] uses the following version of (LL_p^*):

(LL_p^K) $\vdash x = y \supset A \supset [y//x]A$, provided that

    (i) x is free in A and y is free for x in A,

    (ii) y is not free in the scope of a modal operator in A, and

    (iii)   in the scope of any modal operator in A, $[y//x]A$ either replaces all or no occurrences of x by y.

Evidently, this is a lot more restrictive than (LL_p^*). For example, (LL_p^*) allows

$$\vdash x = y \supset \Box Gxy \supset \Box Gyy \quad \text{and}$$
$$\vdash x = y \supset (Fx \lor \Diamond Gxy) \supset (Fy \lor \Diamond Gxy),$$

which can't be derived in Kutz's system (which is therefore incomplete).)

Lemma 7.15 (Leibniz' Law with sequences).
For any $\mathfrak{L}$-formula A and variables $x_1, \dots, x_n, y_1, \dots, y_n$ such that the $x_1, \dots, x_n$ are pairwise distinct,

  (LL_n^*) $\vdash_L x_1 = y_1 \land \dots \land x_n = y_n \supset A \supset [y_1, \dots, y_n/x_1, \dots, x_n]A$, *provided each y_i is really free for x_i in $[y_1, \dots, y_{i-1}/x_1, \dots, x_{n-1}]A$.*

For $i = 1$, the proviso is meant to say that y_1 is really free for x_1 in A.

Proof. By induction on n. For $n = 1$, (LL_n^*) is (LL*). Assume then that $n > 1$ and that each y_i in $y_1, \dots, y_n$ is really free for x_i in $[y_1, \dots, y_{i-1}/x_1, \dots, x_{n-1}]A$. Let z be some variable not in A, $x_1, \dots, x_n, y_1, \dots, y_n$. So z is really free for x_n in A. By (LL*),

$$\vdash_L x_n = z \supset A \supset [z/x_n]A. \tag{1}$$

By induction hypothesis,

$$\vdash_L x_1 = y_1 \land \dots \land x_{n-1} = y_{n-1} \supset$$
$$[z/x_n]A \supset [y_1, \dots, y_{n-1}/x_1, \dots, x_{n-1}][z/x_n]A. \tag{2}$$

By assumption, y_n is really free for x_n in $[y_1, \dots, y_{n-1}/x_1, \dots, x_{n-1}]A$. Then y_n

is also really free for z in $[y_1, \ldots, y_{n-1}/x_1, \ldots, x_{n-1}][z/x_n]A$. So by (LL*),

$$\vdash_L z = y_n \supset [y_1, \ldots, y_{n-1}/x_1, \ldots, x_{n-1}][z/x_n]A \supset$$
$$[y_n/z][y_1, \ldots, y_{n-1}/x_1, \ldots, x_{n-1}][z/x_n]A. \quad (3)$$

But $[y_n/z][y_1, \ldots, y_{n-1}/x_1, \ldots, x_{n-1}][z/x_n]A$ is $[y_1, \ldots, y_n/x_1, \ldots, x_n]A$. Combining (1)–(3), we therefore have

$$\vdash_L x_1 = y_1 \wedge \ldots \wedge x_{n-1} = y_{n-1} \supset$$
$$x_n = z \wedge z = y_n \supset A \supset [y_1, \ldots, y_n/x_1, \ldots, x_n]A. \quad (4)$$

So by (Sub*),

$$\vdash_L x_1 = y_1 \wedge \ldots \wedge x_{n-1} = y_{n-1} \supset$$
$$x_n = x_n \wedge x_n = y_n \supset A \supset [y_1, \ldots, y_n/x_1, \ldots, x_n]A. \quad (5)$$

Since $\vdash_L x_n = y_n \supset x_n = x_n$ (by either (=R) or (Neg) and ($\forall$ = R)), it follows that

$$\vdash_L x_1 = y_1 \wedge \ldots \wedge x_n = y_n \supset A \supset [y_1, \ldots, y_n/x_1, \ldots, x_n]A. \quad (6)$$

Lemma 7.16 (Cross-substitution).
For any $\mathfrak{L}$-formula A and variables x, y,

(CS) $\vdash_L x = y \supset \Box A \supset \Box(y = z \supset [z/x]A)$, *provided z is not free in A.*

More generally, for any variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ such that the $x_1, \ldots, x_n$ are pairwise distinct,

(CS$_n$) $\vdash_L x_1 = y_1 \wedge \ldots \wedge x_n = y_n \supset \Box A \supset \Box(y_1 = z_1 \wedge \ldots \wedge y_n = z_n \supset [z_1, \ldots, z_n/x_1, \ldots, x_n]A)$, *provided none of $z_1, \ldots, z_n$ is free in A.*

Proof. For (CS), assume z is not free in A. Then

1. $\vdash_L x = z \supset A \supset [z/x]A.$ (LL*)
2. $\vdash_L A \supset (x = z \supset [z/x]A).$ (1)
3. $\vdash_L \Box A \supset \Box(x = z \supset [z/x]A).$ (2, (Nec), (K))
4. $\vdash_L x = y \supset \Box(x = z \supset [z/x]A) \supset \Box(y = z \supset [z/x]A).$ (LL*)
5. $\vdash_L x = y \supset \Box A \supset \Box(y = z \supset [z/x]A).$ (3, 4)

Step 4 is justified by the fact that x is not free in $[z/x]A$ and so x and y are modally separated in $x = z \supset [z/x]A$.

The proof for (CS$_n$) is analogous. Assume none of $z_1, \dots, z_n$ is free in A. Then

1. $\vdash_L x_1 = z_1 \wedge \dots \wedge x_n = z_n \supset A \supset [z_1, \dots, z_n/x_1, \dots, x_n]A.$ (LL$_n^*$)
2. $\vdash_L A \supset (x_1 = z_1 \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A).$ (1)
3. $\vdash_L \Box A \supset \Box(x_1 = z_1 \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A).$ (2, (Nec), (K))
4. $\vdash_L x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset$
   $\Box(x_1 = z_1 \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A) \supset$
   $\Box(y_1 = z_1 \wedge \dots \wedge y_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A).$ (LL$_n^*$)
5. $\vdash_L x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \Box A \supset$
   $\Box(x_1 = z_1 \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A).$ (3, 4)

Step 4 is justified by the fact that none of $x_1, \dots, x_n$ is free in $[z_1, \dots, z_n/x_1, \dots, x_n]A$, and each y_i is really free for x_i in $[y_1, \dots, y_{i-1}/x_1, \dots, x_{i-1}]\Box(x_1 = z_1 \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A)$, i.e. in $\Box(y_1 = z_1 \wedge \dots \wedge y_{i-1} = z_{i-1} \wedge x_i = z_i \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A)$, because x_i and y_i are modally separated in $y_1 = z_1 \wedge \dots \wedge y_{i-1} = z_{i-1} \wedge x_i = z_i \wedge \dots \wedge x_n = z_n \supset [z_1, \dots, z_n/x_1, \dots, x_n]A$.

Lemma 7.17 (Substitution-free Universal Instantiation).
For any $\mathfrak{L}$-formula A and variables x, y,

(FUI**) $\vdash_L \forall x A \supset (E! y \supset \exists x (x = y \wedge A)).$

Proof. Let z be a variable not in $\mathrm{Var}(A), x, y$.

1. $\vdash_L z = y \supset E!y \supset E!z$ (LL*)
2. $\vdash_L \forall x A \supset E!z \supset [z/x]A$ ((FUI*), $z \notin \mathrm{Var}(A)$)
3. $\vdash_L \forall x A \wedge E!y \supset z = y \supset [z/x]A$ (1, 2)
4. $\vdash_L \forall x (x = z \supset \neg A) \supset E!z \supset (z = z \supset [z/x]\neg A)$ ((FUI*), $z \notin \mathrm{Var}(A)$)
5. $\vdash_L E!z \supset z = z$ (($=$R), or ($\forall = $R), (FUI*))
6. $\vdash_L \forall x (x = z \supset \neg A) \supset E!z \supset [z/x]\neg A$ (4, 5)
7. $\vdash_L E!z \supset [z/x]A \supset \exists x (x = z \wedge A)$ (6)
8. $\vdash_L \forall x A \wedge E!y \supset z = y \supset \exists x (x = z \wedge A)$ (1, 3, 7)
9. $\vdash_L z = y \supset \exists x (x = z \wedge A) \supset \exists x (x = y \wedge A)$ ((LL*), $z \notin \mathrm{Var}(A)$)
10. $\vdash_L \forall x A \wedge E!y \supset z = y \supset \exists x (x = y \wedge A)$ (8, 9)
11. $\vdash_L \forall z (\forall x A \wedge E!y) \supset \forall z (z = y \supset \exists x (x = y \wedge A))$ (10, (UG), (UD))
12. $\vdash_L \forall x A \wedge E!y \supset \forall z (z = y \supset \exists x (x = y \wedge A))$ (11, (VQ))
13. $\vdash_L \forall z (z = y \supset \exists x (x = y \wedge A)) \supset y = y \supset \exists x (x = y \wedge A)$ ((FUI*), $z \notin \mathrm{Var}(A)$)
14. $\vdash_L E!y \supset y = y$ (($=$R), or ($\forall = $R), (FUI*))
15. $\vdash_L \forall z (z = y \supset \exists x (x = y \wedge A)) \supset E!y \supset \exists x (x = y \wedge A)$ (13, 14)
16. $\vdash_L \forall x A \supset E!y \supset \exists x (x = y \wedge A)$ (12, 15)

(FUI*) can also be derived from (FUI**), so we could just as well have used (FUI**) as axiom instead of (FUI*).

8 Object-language substitution

8.1 The substitution operator

The failure of the substitution lemma in non-functional counterpart semantics indicates an expressive defect of the language interpreted by that semantics. Return to Leibniz' Law. If $x = y$ then whatever is true of x is true of y. Yet

$$x = y \supset \Diamond Gxy \supset \Diamond Gyy$$

is invalid in the class of non-functional counterpart structures. The problem is that $\Diamond Gyy$ does not "say about y" what $\Diamond Gxy$ "says about x". Roughly speaking, $\Diamond Gxy$ says that at some accessible world, some counterpart of x (and therefore y) is G-related to some (possibly different) counterpart of y (and therefore x), but $\Diamond Gyy$ says that some counterpart of y (and therefore x) is G-related to itself. Let's briefly look at how this problem arises, and how it can be solved.

Modal operators, we assume, shift the point of evaluation. In counterpart semantics, when the point of evaluation is shifted from w to w', the semantic value of every individual constant and variable shifts to the counterpart of the previous value, following some counterpart relation C. If an individual at w has no counterpart at w', the relevant terms become empty. If an individual has multiple C-counterparts, we may think of the corresponding terms as becoming "ambiguous", denoting all the counterparts at the same time. To verify $\Box Fx$, we require that Fx is true at all accessible worlds under all "disambiguations".

An important question now is whether these disambiguations are uniform or mixed: should $\Box Gxx$ be true iff at all accessible worlds (relative to all counterpart relations), all x counterparts are G-related to *themselves* (uniform) or to *one another* (mixed)? On the mixed account, $\Box x = x$ becomes invalid, as does $\Box(Fx \lor \neg Fx)$, even if (the individual denoted by) x exists at all worlds. The semantics also becomes more complicated because a mixed disambiguation cannot be represented by a standard

assignment function. If we say that $\Box A$ is true under an assignment g iff A is true at all accessible worlds under all assignments g' suitably related to g, and assignments are functions (rather than relations), we automatically get uniform disambiguations. In definition 2.9, we therefore have opted for a uniform disambiguation. This may not be the best choice for every application, but it is the only one we're going to explore.

The present issue might remind you of the old observation that a sentence like 'Brutus killed himself' can be understood either as an application of a *monadic* predicate 'killing himself' to the subject Brutus, or as an application of the binary 'killing' to Brutus and Brutus. Peter Geach once suggested a syntactic mechanism for distinguishing these readings, introducing an operator $\langle z : x, y \rangle$ that turns a binary expression into a unary expression: while Gxy is satisfied by pairs of individuals, $\langle z : x, y \rangle Gxy$ is satisfied by a single individual. The operator $\langle z : x, y \rangle$, which might be read 'z is an x and a y such that', acts as a quantifier that binds both x and y.

A similar trick can be used in our modal context. On our uniform reading, $\Box x = x$ says (roughly) that all counterparts of x are self-identical at all accessible worlds. To say that at all accessible worlds (and under all counterpart relations), all x counterparts are identical to all x counterparts, we could instead say $\langle x : y, z \rangle \Box y = z$. The effect of $\langle x : y, z \rangle$ is to introduce two variables y and z that co-refer with x. By using distinct but co-refering variables in a modal context, we can express relations between possibly distinct counterparts; by using the same variable, we make sure that the same counterpart must be assigned to every occurrence.

With $\langle x : y, z \rangle \Box y = z$, we actually end up with *three* co-referring variables: y and z are made to co-refer with x, but we also have x itself. The job can also be done with $\langle x : y \rangle \Box x = y$ – read: 'x is a y such that …'.

To see the use of this operator, consider the following two sentences, which look at first glance like simple applications of universal instantiation.

$$\forall x \Box Gxy \supset \Box Gyy; \tag{1}$$

$$\forall x \Diamond Gxy \supset \Diamond Gyy. \tag{2}$$

Suppose for a moment that we have at most one counterpart relation from any world to another, so that we can ignore the quantification over counterpart relations. The first formula then says that if all things x are such that all x counterparts are G-related to all y counterparts, then all y counterparts are G-related to themselves. In

a total counterpart model, this must be true. (2), however, is not valid in the class of total structures. If all things x are such that some x counterpart is G-related to some y counterpart, it only follows that some y counterpart is G-related to some y counterpart; it does not follow that some y counterpart is G-related to itself.

With the two distinct variables x and y, the formula $\Diamond Gxy$ looks at arbitrary combinations of x counterparts and y counterparts, even if the variables co-refer. By contrast, $\Diamond Gyy$ only looks at single y counterparts and checks whether one of them is G-related to itself. To prevent this "capturing" of y in the consequent of (2), we can use the Geach quantifier:

$$\forall x \Diamond Gxy \supset \langle y : x \rangle \Diamond Gxy \tag{2'}$$

Having multiple counterpart relations makes no essential difference to these considerations. $\Diamond Gyy$ is true at w iff Gyy is true at some accessible world under some assignment of a y-counterpart to 'y'. On the other hand, given $x = y$, $\Diamond Gxy$ is true at w iff there is an accessible world at which some counterpart of the pair $\langle x, y \rangle$ ($= \langle x, x \rangle = \langle y, y \rangle$) satisfies Gxy.

The situation is analogous for Leibniz' Law. As noted above, (4) is invalid, while (3) is valid.

$$x = y \supset \Box Gxy \supset \Box Gyy; \tag{3}$$

$$x = y \supset \Diamond Gxy \supset \Diamond Gyy. \tag{4}$$

In (4), the substituted variable y again gets captured by the other occurrence of y in the scope of the diamond. $\Diamond Gyy$ does not say of y what Gxy says of x. We know from proposition 7.3 that the standard language of modal predicate logic does not provide a general way to say of y what some formula A says of x.

The Geach quantifier provides the missing resource. In place of (4), we could write

$$x = y \supset \Diamond Gxy \supset \langle y : x \rangle \Diamond Gxy \tag{4'}$$

Here, as above, the Geach quantifier $\langle y : x \rangle$ functions as an object-language substitution operator.

Some of the mysterious features of non-functional counterpart logics become more readily intelligible when we have a substitution operator. A perhaps more

familiar tool that would also do the job is lambda abstraction, introduced to modal logic in [Carnap 1947] and [Stalnaker and Thomason 1968]. With lambda abstraction, we would write $(\lambda x.A)y$ instead of $\langle y : x \rangle A$ to express that y is an x such that A. For our purposes, the substitution operator will prove more perspicuous.

> **Definition 8.1 (The language of modal predicate logic with substitution).**
> The *language of modal predicate logic with substitution* is the standard language of modal predicate logic from definition 2.3 with an added construct $\langle : \rangle$ and the rule that whenever x, y are variables and A is a formula, then $\langle y : x \rangle A$ is a formula.

As for the semantics: just as $\forall x A$ is true relative to an assignment g iff A is true relative to all x-variants of g (on the relevant domain), $\langle y : x \rangle A$ is true relative to g iff A is true relative to the x-variant of g that maps x to $g(y)$. In our modal framework:

> **Definition 8.2 (Semantics for the substitution operator).**
> For any counterpart model $\mathfrak{M}$, any world w in $\mathfrak{M}$, assignment g on U_w, formula A, and variables x, y,
>
> $$\mathfrak{M}, w, g \Vdash \langle y : x \rangle A \ \text{ iff } \ \mathfrak{M}, w, g \circ [y/x] \Vdash A.$$

Substitution operators have great expressive power. As [Kuhn 1980] shows (in effect), if a language has substitution operators, it no longer needs variables or individual constants in its atomic formulas: instead of Fx, we can simply say F, with the convention that the implicit variable is always x (for binary predicates, the first variable is x, the second y, etc.); Fy turns into $\langle y : x \rangle F$, Gyz into $\langle y : x \rangle \langle z : y \rangle G$. Similarly, $\forall x Fx$ can be replaced by $\forall F$, and $\forall y Gxy$ by $\forall \langle y : z \rangle \langle x : y \rangle \langle z : x \rangle G$. So we also don't need different quantifiers for different variables. I will not exploit the full power of substitution operators – mainly for the sake of familiarity. Our languages with substitution operators still have ordinary formulas $Px_1 \ldots x_n$ and quantifiers $\forall x, \forall y$, etc.

It can sometimes be useful to have polyadic quantifiers like $\langle y_1, y_2 : x_1, x_2 \rangle$, which says '$y_1$ is an x_1 and y_2 an x_2 such that', where

$$\mathfrak{M}, w, g \Vdash \langle y_1, y_2 : x_1, x_2 \rangle A \ \text{ iff } \ \mathfrak{M}, w, g \circ [y_1, y_2/x_1, x_2] \Vdash A.$$

Geach's $\langle x : y, z \rangle$ is then equivalent to $\langle x, x : y, z \rangle$.

We don't need to introduce polyadic substitution operators as primitive, however. We can instead treat them as abbreviations for iterated monadic substitution.

We can't simply define $\langle y_1, y_2 : x_1, x_2 \rangle A$ as $\langle y_1 : x_1 \rangle \langle y_2 : x_2 \rangle A$, since the bound variable x_1 might capture y_2, as in the swapping operator $\langle x, y : y, x \rangle$. We have to store the original value of y_2 in a temporary variable z: $\langle y_2 : z \rangle \langle y_1 : x_1 \rangle \langle z : x_2 \rangle$.

Definition 8.3 (Substitution sequences).
For any $n > 1$, formula A and variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ in which the $x_1, \ldots, x_n$ are all distinct, let

$$\langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A$$

abbreviate
$$\langle y_n : z \rangle \langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle \langle z : x_n \rangle A,$$

where z is the alphabetically first variable not in A or $x_1, \ldots, x_n$.

Lemma 8.1 (Substitution sequence semantics).
For any counterpart model $\mathfrak{M}$, any world w in $\mathfrak{M}$, assignment g on U_w, formula A, and variables $x_1, \ldots, x_n, y_1, \ldots, y_n$, where the $x_1, \ldots, x_n$ are all distinct,

$$\mathfrak{M}, w, g \Vdash \langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A \text{ iff } \mathfrak{M}, w, g \circ [y_1, \ldots, y_n / x_1, \ldots, x_n] \Vdash A.$$

Proof. The proof is by induction on n.

By definition 8.3, $\mathfrak{M}, w, g \Vdash \langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A$ iff $\mathfrak{M}, w, g \Vdash \langle y_n : z \rangle$ $\langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle \langle z : x_n \rangle A$, for some z not in A or $x_1, \ldots, x_{n-1}$. By definition 8.2, $\mathfrak{M}, w, g \Vdash \langle y_n : z \rangle \langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle \langle z : x_n \rangle A$ iff $\mathfrak{M}, w, g \circ$ $[y_n / z] \Vdash \langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle \langle z : x_n \rangle A$. By induction hypothesis, the latter holds iff $\mathfrak{M}, w, g \circ [y_n / z] \circ [y_1, \ldots, y_{n-1} / x_1, \ldots, x_{n-1}] \Vdash \langle z : x_n \rangle A$, that is, iff $\mathfrak{M}, w, g \circ [y_n / z] \circ [y_1, \ldots, y_{n-1} / x_1, \ldots, x_{n-1}] \circ [z / x_n] \Vdash A$ by definition 8.2.

Now $[y_n/z] \circ [y_1, \ldots, y_{n-1}/x_1, \ldots, x_{n-1}] \circ [z/x_n]$ is the substitution σ such that

$$\sigma(x) = [y_n/z]([y_1, \ldots, y_{n-1}/x_1, \ldots, x_{n-1}]([z/x_n](x))).$$

Since $z \notin x_1, \ldots, x_{n-1}$, this means that

$$\sigma(x_n) = y_n,$$
$$\sigma(x_i) = y_i \text{ for } x_i \in \{x_1, \ldots, x_{n-1}\},$$
$$\sigma(z) = y_n,$$

and $\sigma(x) = x$ for every other variable x. Since $z \notin \mathrm{Var}(A)$, it follows by the locality lemma (which is easily extended to languages with object-language substitution) that $\mathfrak{M}, w, g \circ \sigma \Vdash A$ iff $\mathfrak{M}, w, g \circ [y_1, \ldots, y_n/x_1, \ldots, x_n] \Vdash A$.

8.2 Substitution logics

With an object-language substitution operator, we can express the substitution principles (FUI), (LL), and (Sub) without any restrictions on the relevant terms:

(FUIˢ) $\forall xA \supset (Ey \supset \langle y : x \rangle A)$,

(LLˢ) $x = y \supset (A \supset \langle y : x \rangle A)$,

(Subˢ) if $\vdash_L A$, then $\vdash_L \langle y : x \rangle A$.

To get a complete logic, we also need some rules for the substitution operator. An obvious suggestion would be the lambda-conversion principle

$$\langle y : x \rangle A \leftrightarrow [y/x]A.$$

This would allow us to move back and forth between, for example, $\langle y : x \rangle Fx$ and Fy. But we've seen that these transitions are not always sound: the move from $\langle y : x \rangle A$ to $[y/x]A$ requires that y is really free for x in A; the other direction requires that y and x are modally separated in A. (See lemmas 7.1 and 7.2.) We are going to have the following, more restricted principles:

(SC1) $\langle y : x \rangle A \leftrightarrow [y/x]A$, provided y and x are modally separated in A.

(SC2) $\langle y : x \rangle A \supset [y/x]A$, provided y is really free for x in A.

I haven't yet explained what $[y/x]A$ is. We could use a naive definition. Then we would have to add the restriction that no free occurrence of x in A lies in the scope of a quantifier that binds y. For continuity with the earlier chapters, I will instead extend definition 3.1 so that it can be applied to formulas that involve the substitution operator.

Definition 8.4 (Substitution).

A *substitution* (on a set of variables Var) is a total function $\sigma : \text{Var} \to \text{Var}$. Application of a substitution σ to a formula A is defined as follows.

$\sigma(Px_1 \ldots x_n) = P\sigma(x_1) \ldots \sigma(x_n)$

$\sigma(\neg A) = \neg \sigma(A)$

$\sigma(A \supset B) = \sigma(A) \supset \sigma(B)$

$$\sigma(\forall z A) = \begin{cases} \forall v \sigma^{v \mapsto v}([v/z]A) & \text{if there is an } x \in \text{FV}(\forall z A) \text{ with } \sigma(x) = \sigma(z) \\ \forall \sigma(z) \sigma(A) & \text{otherwise,} \end{cases}$$
where v is the alphabetically first variable not in $\text{FV}(\sigma(A)) \cup \text{FV}(A)$ and $\sigma^{v \mapsto v}$ is the substitution that maps v to v and otherwise coincides with σ.

$$\sigma(\langle y : z \rangle A) = \begin{cases} \langle \sigma(y) : z \rangle \sigma^{v \mapsto v}([v/z]A) & \text{if there is an } x \in \text{FV}(\langle y : z \rangle A) \text{ with } \sigma(x) = \sigma(z) \\ \langle \sigma(y) : \sigma(z) \rangle \sigma(A) & \text{otherwise,} \end{cases}$$
where v is the alphabetically first variable not in $\text{FV}(\sigma(A)) \cup \text{FV}(A)$ and $\sigma^{v \mapsto v}$ is the substitution that maps v to v and otherwise coincides with σ.

$\sigma(\square A) = \square \sigma(A)$.

As before, $[y_1 \ldots y_n / x_1 \ldots x_n]$ is the substitution σ that maps x_i to y_i (for $1 \leq i \leq n$) and every other variable to itself.

The clause for the substitution operator is exactly parallel to the one for the universal quantifier, and the underlying motivation is the same. For example, $[y/x]\langle y_2 : y \rangle x \neq y$ is $\langle y_2 : z \rangle y \neq z$, rather than $\langle y_2 : y \rangle y \neq y$.

(SC1) and (SC2) are not enough. We need further principles telling us how $\langle y : x \rangle$ behaves when y is not really free for x in A. For example, $\langle y : x \rangle \neg A$ should always

entail $\neg\langle y : x\rangle A$, even if y is not really free for x in A. More generally, the substitution operator commutes with every non-modal operator as long as there is no clash of bound variables:

(S¬) $\langle y : x\rangle\neg A \leftrightarrow \neg\langle y : x\rangle A$,

(S⊃) $\langle y : x\rangle(A \supset B) \leftrightarrow (\langle y : x\rangle A \supset \langle y : x\rangle B)$,

(S∀) $\langle y : x\rangle\forall z A \leftrightarrow \forall z\langle y : x\rangle A$, provided $z \notin \{x, y\}$,

(SS1) $\langle y : x\rangle\langle y_2 : z\rangle A \leftrightarrow \langle y_2 : z\rangle\langle y : x\rangle A$, provided $z \notin \{x, y\}$ and $y_2 \neq x$.

Substitution does not commute with the box. Roughly speaking, this is because $\langle y : x\rangle\Box A(x, y)$ says that at all accessible worlds, all counterparts x' and y' of y are $A(x', y')$, while $\Box\langle y : x\rangle A(x, y)$ says that at all accessible worlds, every counterpart $x' = y'$ of y is such that $A(x', y')$. In the first case, x' and y' may be different counterparts of y, while in the second case, they must be the same. Thus $\langle y : x\rangle\Box A$ entails $\Box\langle y : x\rangle A$, but the other direction holds only if either y does not have multiple counterparts at accessible worlds (relative to the same counterpart relation), or at most one of x and y occurs freely in A (including the special case where x and y are the same variable). We have:

(S□) $\langle y : x\rangle\Box A \supset \Box\langle y : x\rangle A$,

(S◇) $\langle y : x\rangle\Diamond A \supset \Diamond\langle y : x\rangle A$, provided at most one of x, y is free in A.

These principles largely make (SC1) and (SC2) redundant. We only need to add the special case for substituting free variables in atomic formulas and in substitution operators, as well as a principle for vacuous substitutions:

(SAt) $\langle y : x\rangle Px_1 \ldots x_n \leftrightarrow P[y/x]x_1 \ldots [y/x]x_n$.

(SS2) $\langle y : x\rangle\langle x : z\rangle A \leftrightarrow \langle y : z\rangle\langle y : x\rangle A$.

(VS) $A \leftrightarrow \langle y : x\rangle A$, provided x is not free in A.

Let FK^s be the smallest system that contains all $\mathfrak{L}_s$-instances of the substitution axioms (S¬), (S⊃), (S∀), (SS1), (S□), (S◇), as well as (Taut), (UD), (VQ), (FUI^s), (∀Ex), (=R), (LL^s), (K), and that is closed under (MP), (UG), (Nec) and (Sub^s).

Let NK^s be the smallest system that contains all $\mathfrak{L}_s$-instances of the substitution axioms (S¬), (S⊃), (S∀), (SS1), (S□), (S◇), as well as (Taut), (UD), (VQ), (Neg), (NA), (∀=R), (K), (FUI^s), (LL^s), and that is closed under (MP), (UG), (Nec) and (Sub^s).

> **Lemma 8.2 (Soundness of the substitution axioms).**
> *Every instance of* (S¬), (S⊃), (S∀), (SS1), (S□), (S◊), (SAt), (SS2), *and* (VS) *is valid on every counterpart structure.*

Proof.

1. (S¬). $\mathfrak{M}, w, g \Vdash \langle y : x \rangle \neg A$ iff $\mathfrak{M}, w, g^{[y/x]} \Vdash \neg A$ by definition 8.2, iff $\mathfrak{M}, w, g^{[y/x]} \not\Vdash A$ by definition 2.9, iff $\mathfrak{M}, w, g \not\Vdash \langle y : x \rangle A$ by definition 8.2, iff $\mathfrak{M}, w, g \Vdash \neg \langle y : x \rangle A$ by definition 2.9.

2. (S⊃). $\mathfrak{M}, w, g \Vdash \langle y : x \rangle (A \supset B)$ iff $\mathfrak{M}, w, g^{[y/x]} \Vdash A \supset B$ by definition 8.2, iff $\mathfrak{M}, w, g^{[y/x]} \not\Vdash A$ or $\mathfrak{M}, w, g^{[y/x]} \Vdash B$ by definition 2.9, iff $\mathfrak{M}, w, g \not\Vdash \langle y : x \rangle A$ or $\mathfrak{M}, w, g \Vdash \langle y : x \rangle B$ by definition 8.2, iff $\mathfrak{M}, w, g \Vdash \langle y : x \rangle A \supset \langle y : x \rangle B$ by definition 2.9.

3. (S∀). Assume $z \notin \{x, y\}$. $\mathfrak{M}, w, g \Vdash \langle y : x \rangle \forall z A$ iff $\mathfrak{M}, w, g^{[y/x]} \Vdash \forall z A$ by definition 8.2, iff $\mathfrak{M}, w, g^{[y/x]^{z \mapsto d}} \Vdash A$ for all $d \in D_w$ by definition 2.9, iff $\mathfrak{M}, w, g^{z \mapsto d} \circ [y/x]$ for all $d \in D_w$ because $z \notin \{x, y\}$, iff $\mathfrak{M}, w, g^{z \mapsto d} \Vdash \langle y : x \rangle A$ for all $d \in D_w$ by definition 8.2, iff $\mathfrak{M}, w, g \Vdash \forall z \langle y : x \rangle A$ by definition 2.9.

4. (SS1). Assume $z \notin \{x, y\}$ and $y_2 \neq x$. Then the function $[y/x] \circ [y_2/z]$ is identical to the function $[y_2/z] \circ [y/x]$. So $\mathfrak{M}, w, g \Vdash \langle y : x \rangle \langle y_2 : z \rangle A$ iff $\mathfrak{M}, w, g^{[y/x] \circ [y_2/z]} \Vdash A$ by definition 8.2, iff $\mathfrak{M}, w, g^{[y_2/z] \circ [y/x]} \Vdash A$, iff $\mathfrak{M}, w, g \Vdash \langle y_2 : z \rangle \langle y : x \rangle A$ by definition 8.2.

5. (S□). Assume $\mathfrak{M}, w, g \not\Vdash \Box \langle y : x \rangle A$. By definitions 2.9 and 8.2, this means that $\mathfrak{M}, w', g'^{[y/x]} \not\Vdash A$ for some w', g' such that $w, g \rhd w', g'$; that is, for some w', g' such that wRw' and there is a $C \in K_{w,w'}$ so that g' assigns to every variable z a C-counterpart of its g value (or nothing if there is no such counterpart). This means that for all z, $g'^{[y/x]}(z)$ is a C-counterpart of $g^{[y/x]}(z)$ (or undefined if there is none), since $g'^{[y/x]}(x) = g'(y)$ is a C-counterpart of $g(y) = g^{[y/x]}(x)$ (or undefined if there is none). So $w, g^{[y/x]} \rhd w', g'^{[y/x]}$. And so $w', g^* \not\Vdash A$ for some w', g^* such that $w, g^{[y/x]} \rhd w, g^*$. So $\mathfrak{M}, w, g \Vdash \langle y : x \rangle \Box A$ by definitions 2.9 and 8.2.

6. (S$\Diamond$). Assume $\mathfrak{M}, w, g \Vdash \langle y : x \rangle \Diamond A$ and at most one of x, y is free in A. By definitions 2.9 and 8.2, $w', g^* \Vdash A$ for some w', g^* such that wRw' and there is a $C \in K_{w,w'}$ for which g^* assigns to every variable z a C-counterpart of its g value (or nothing if there is none). We have to show that there is a w'-image g' of g at w such that $\mathfrak{M}, w, g'^{[y/x]} \Vdash A$, since then $\mathfrak{M}, w, g \Vdash \Diamond \langle y : x \rangle A$.

   If x is the same variable as y, then $g^*(x) = g^*(y)$ is a C-counterpart at w' of $g^{[y/x]}(x) = g^{[y/x]}(y) = g(x) = V(y)$ at w (or undefined if there is none), so we can choose g^* itself as g'. We then have $\mathfrak{M}, w, g'^{[y/x]} \Vdash A$ because $g'^{[y/x]} = g'$.

   Else if x is not free in A, let g' be some x-variant of g^* at w' such that $g^*(x)$ is some C-counterpart at w' of $g(x)$ at w (or undefined if there is none). Since $g^*(y)$ is a C-counterpart at w' of $g^{[y/x]}(y) = V(y)$ at w (or undefined if there is none), g' is a w'-image of g at w. Moreover, $g'^{[y/x]}$ and g^* differ at most in the value of x; by the locality lemma, it follows that $\mathfrak{M}, w', g'^{[y/x]} \Vdash A$.

   Else if y is not free in A, let g' be like g^* except that $g'(y) = g^*(x)$ and $g'(x)$ is some C-counterpart at w' of $g(x)$ at w (or undefined if there is none). Since $g'(y) = g^*(x)$ is a C-counterpart at w' of $g^{[y/x]}(x) = g(y)$ at w (or undefined if there is none), g' is a w'-image of g at w. Moreover, $g'^{[y/x]}$ and g^* differ at most in the value of y; in particular, $g'^{[y/x]}(x) = g'(y) = g^*(x)$. By the locality lemma, it follows that $\mathfrak{M}, w', g'^{[y/x]} \Vdash A$.

7. (SAt). $\mathfrak{M}, w, g \Vdash \langle y : x \rangle Px_1 \ldots x_n$ iff $\mathfrak{M}, w, g^{[y/x]} \Vdash Px_1 \ldots x_n$ by definition 8.2, iff $\mathfrak{M}, w, g \Vdash [y/x]Px_1 \ldots x_n$ by lemma 7.2.

8. (SS2). $\mathfrak{M}, w, g \Vdash \langle y : x \rangle \langle x : z \rangle A$ iff $\mathfrak{M}, w, g^{[y/x] \circ [x/z]} \Vdash A$ by definition 8.2, iff $\mathfrak{M}, w, g^{[y/z] \circ [y/x]} \Vdash A$ because $[y/x] \circ [x/z] = [y/z] \circ [y/x]$, iff $\mathfrak{M}, w, g \Vdash \langle y : z \rangle \langle y : x \rangle A$ by definition 8.2.

9. (VS). By definition 8.2, $\mathfrak{M}, w, g \Vdash \langle y : x \rangle A$ iff $\mathfrak{M}, w, g^{[y/x]} \Vdash A$. If x is not free in A then $g^{[y/x]}$ coincides with g at w for all free variables in A. By the locality lemma, it follows that $\mathfrak{M}, w, g^{[y/x]} \Vdash A$ iff $\mathfrak{M}, w, g \Vdash A$. So $\mathfrak{M}, w, g \Vdash \langle y : x \rangle A$ iff $\mathfrak{M}, w, g \Vdash A$.

Lemma 8.3 (Soundness of (FUI^s)**,** (LL^s)**, and** (Sub^s)**).**
Every instance of (FUI^s) *and* (LL^s) *is valid in every counterpart structure. If all elements of a set of formulas* Γ *are valid on a counterpart structure* $\mathfrak{S}$*, and* Γ *is extended by* (Sub^s)*, then the new sentences are still valid on* $\mathfrak{S}$*.*

Proof.

1. (FUI^s). Assume $\mathfrak{M}, w, g \Vdash \forall x A$ and $\mathfrak{M}, w, g \Vdash Ey$. By lemma 3.2 and definition 2.9, the latter means that $g(y) \in D_w$, while the former means that $\mathfrak{M}, w, g^{x \mapsto d} \Vdash A$ for all $d \in D_w$. So in particular, $\mathfrak{M}, w, g^{x \mapsto g(y)} \Vdash A$. And so $\mathfrak{M}, w, g \Vdash \langle y : x \rangle A$ by definition 8.2.

2. (LL^s). Assume $\mathfrak{M}, w, g \Vdash x = y$ and $\mathfrak{M}, w, g \Vdash A$. By definition 2.9, we have $g(x) = g(y)$, and so $\mathfrak{M}, w, g \Vdash \langle y : x \rangle A$ by definition 8.2.

3. (Sub^s). Assume for contraposition that $\mathfrak{M}, w, g \nVdash \langle y : x \rangle A$. By definition 8.2, this means that $\mathfrak{M}, w, g' \nVdash A$, where g' is the x-variant of g with $g'(x) = g(y)$. So A is not valid on every counterpart structure.

Theorem 8.4 (Soundness of FK^s**).**
Every member of FK^s *is valid on every total counterpart structure.*

Proof. Immediate from lemmas 7.4, 7.5. 8.2, and 8.3.

Theorem 8.5 (Soundness of NK^s**).**
Every member of NK^s *is valid on every single-domain counterpart structure.*

Proof. Immediate from lemmas 7.9, 7.5, 8.2, and 8.3.

8.3 Some consequences

As in earlier sections, we'll prove some consequences of the above axiomatisations. As before, these will hold for all logics that extend FKˢ or NKˢ by further axioms.

Definition 8.5 (Positive logics).
A set of $\mathfrak{L}_s$-sentences is a *positive logic* if it includes FKˢ and is closed under (MP), (UG), (Nec) and (Subˢ).

Definition 8.6 (Negative logics).
A set of $\mathfrak{L}_s$-sentences is a *negative logic* if it includes NKˢ and is closed under (MP), (UG), (Nec) and (Subˢ).

Until further notice, let L be any positive or negative logic.

Lemmas 3.10 and 3.11 can be established as in section 3.4. Let's turn right away to some derived principles about substitution.

Lemma 8.6 (Substitution expansion).
If A is an $\mathfrak{L}$-formula and x, y, z $\mathfrak{L}$-variables, then

(SE1) $\vdash_L A \leftrightarrow \langle x : x \rangle A$;

(SE2) $\vdash_L \langle y : x \rangle A \leftrightarrow \langle y : z \rangle \langle z : x \rangle A$, *provided z is not free in A.*

Proof. We first prove (SE1), by induction on A.

1. A is atomic. Then $\vdash_L \langle x : x \rangle A \leftrightarrow [x/x]A$ by (SAt), and so $\vdash_L \langle x : x \rangle A \leftrightarrow A$ because $[x/x]A = A$.

2. A is $\neg B$. By induction hypothesis, $\vdash_L B \leftrightarrow \langle x : x \rangle B$. So by (PC), $\vdash_L \neg B \leftrightarrow \neg \langle x : x \rangle B$. And by $\langle S\neg \rangle$, $\vdash_L \langle x : x \rangle \neg B \leftrightarrow \neg \langle x : x \rangle B$.

3. A is $B \supset C$. By induction hypothesis, $\vdash_L B \leftrightarrow \langle x : x \rangle B$ and $\vdash_L C \leftrightarrow \langle x : x \rangle C$. So $\vdash_L (B \supset C) \leftrightarrow (\langle x : x \rangle B \supset \langle x : x \rangle C)$. And by $\langle S \supset \rangle$, $\vdash_L \langle x : x \rangle (B \supset C) \leftrightarrow (\langle x : x \rangle B \supset \langle x : x \rangle C)$.

4. A is $\forall z B$. If $z = x$ then $\vdash_L \forall x B \leftrightarrow \langle x : x \rangle \forall x B$ by (VS). If $z \neq x$ then by induction hypothesis, $\vdash_L B \leftrightarrow \langle x : x \rangle B$; by (UG) and (UD), $\vdash_L \forall z B \leftrightarrow \forall z \langle x : x \rangle B$; and $\vdash_L \langle x : x \rangle \forall z B \leftrightarrow \forall z \langle x : x \rangle B$ by (S$\forall$).

5. A is $\langle y : z \rangle B$. If $z = x$, then $\vdash_L \langle y : x \rangle B \leftrightarrow \langle x : x \rangle \langle y : x \rangle B$ by (VS). If $z \neq x$ then by induction hypothesis, $\vdash_L B \leftrightarrow \langle x : x \rangle B$; by (Sub^s) and (S$\supset$), $\vdash_L \langle y : z \rangle B \leftrightarrow \langle y : z \rangle \langle x : x \rangle B$; and $\vdash_L \langle x : x \rangle \langle y : z \rangle B \leftrightarrow \langle y : z \rangle \langle x : x \rangle B$ by (SS1) (if $y \neq x$) or (SS2) (if $y = x$).

6. A is $\Box B$. By (S$\Box$), $\vdash_L \langle x : x \rangle \Box B \supset \Box \langle x : x \rangle B$. Conversely, since at most one of x, x is free in $\neg B$, by (S$\Diamond$), $\vdash_L \langle x : x \rangle \Diamond \neg B \supset \Diamond \langle x : x \rangle \neg B$. Contraposing and unraveling the definition of the diamond, we have $\vdash_L \Box \neg \langle x : x \rangle \neg B \supset \neg \langle x : x \rangle \neg \Box \neg \neg B$. Since $\vdash_L \Box \neg \langle x : x \rangle \neg B \leftrightarrow \Box \langle x : x \rangle B$ and $\vdash_L \neg \langle x : x \rangle \neg \Box \neg \neg B \leftrightarrow \langle x : x \rangle B$ (by (S$\neg$), (Sub^s), (S$\supset$), (Nec) and (K)), this means that $\vdash_L \Box \langle x : x \rangle B \supset \langle x : x \rangle \Box B$.

Now (SE2).

By (VQ), $\vdash_L \langle y : x \rangle A \leftrightarrow \langle y : z \rangle \langle y : x \rangle A$. Also, $\vdash_L \langle y : x \rangle \langle y : z \rangle A \leftrightarrow \langle y : z \rangle \langle y : x \rangle A$ by (SS1) (if $y \neq x$) or (SS2) (if $y = x$). Moreover, by (SS2), $\vdash_L \langle y : z \rangle \langle z : x \rangle A \leftrightarrow \langle y : x \rangle \langle y : z \rangle A$. So by (PC), $\vdash_L \langle y : x \rangle A \leftrightarrow \langle y : z \rangle \langle z : x \rangle A$.

Lemma 8.7 (Substituting bound variables).

For any $\mathfrak{L}$-sentence A and variables x, y,

(SBV) $\vdash_L \forall x A \leftrightarrow \forall y \langle y : x \rangle A$, *provided y is not free in A.*

Proof.

$$1. \quad \vdash_L \forall y \langle y : x \rangle A \supset Ex \supset \langle x : y \rangle \langle y : x \rangle A. \qquad (\text{FUI}^s)$$

$$2. \quad \vdash_L \langle x : y \rangle \langle y : x \rangle A \leftrightarrow A. \qquad ((\text{SE1}), (\text{SE2}))$$

$$3. \quad \vdash_L \forall x \forall y \langle y : x \rangle A \supset \forall x Ex \supset \forall x A. \qquad (1, 2, (\text{UG}), (\text{UD}))$$

$$4. \quad \vdash_L \forall x \forall y \langle y : x \rangle A \supset \forall x A. \qquad (3, (\forall Ex))$$

$$5. \quad \vdash_L \forall y \langle y : x \rangle A \supset \forall x \forall y \langle y : x \rangle A. \qquad (\text{VQ})$$

$$6. \quad \vdash_L \forall y \langle y : x \rangle A \supset \forall x A. \qquad (4, 5)$$

$$7. \quad \vdash_L \forall x A \supset Ey \supset \langle y : x \rangle A. \qquad (\text{FUI}^s)$$

$$8. \quad \vdash_L \forall y \forall x A \supset \forall y \langle y : x \rangle A. \qquad (7, (\text{UG}), (\text{UD}), (\forall Ex))$$

$$9. \quad \vdash_L \forall x A \supset \forall y \forall x A. \qquad ((\text{VQ}), y \text{ not free in } A)$$

$$10. \quad \vdash_L \forall x A \supset \forall y \langle y : x \rangle A. \qquad (8, 9)$$

$$11. \quad \vdash_L \forall x A \leftrightarrow \forall y \langle y : x \rangle A. \qquad (6, 10)$$

Lemma 8.8 (Substituting empty variables).
For any $\mathfrak{L}$-sentence A and variables x, y,

$$(\text{SEV}) \quad \vdash_L x \neq x \wedge y \neq y \supset (A \leftrightarrow \langle y : x \rangle A).$$

Proof. (SEV) is trivial if L is positive, in which case $\vdash_L x = x$. For negative L, we proceed by induction on A (in smaller font, to make the lines fit).

1. *A is atomic.* If $x \notin \text{Var}(A)$ then $\vdash_L A \leftrightarrow \langle y : x \rangle A$ by (VS), and so $\vdash_L x \neq x \wedge y \neq y \supset (A \leftrightarrow \langle y : x \rangle A)$ by (PC). If $x \in \text{Var}(A)$ then by (Neg),

$$\vdash_L x \neq x \wedge y \neq y \supset \neg A. \tag{1}$$

Also by (Neg), $\vdash_L x \neq x \wedge y \neq y \supset \neg [y/x]A$. By (SAt), $\vdash_L [y/x]A \leftrightarrow \langle y : x \rangle A$, and so $\vdash_L \neg [y/x]A \leftrightarrow \neg \langle y : x \rangle A$. So

$$\vdash_L x \neq x \wedge y \neq y \supset \neg \langle y : x \rangle A. \tag{2}$$

112

Combining (1) and (2) yields $\vdash_L x \neq x \wedge y \neq y \supset (A \leftrightarrow \langle y : x \rangle A)$.

2. A is $\neg B$. By induction hypothesis, $\vdash_L x \neq x \wedge y \neq y \supset (B \leftrightarrow \langle y : x \rangle B)$. So by (PC), $\vdash_L x \neq x \wedge y \neq y \supset (\neg B \leftrightarrow \neg \langle y : x \rangle B)$, and by (S¬), $\vdash_L x \neq x \wedge y \neq y \supset (\neg B \leftrightarrow \langle y : x \rangle \neg B)$.

3. A is $B \supset C$. By induction hypothesis, $\vdash_L x \neq x \wedge y \neq y \supset (B \leftrightarrow \langle y : x \rangle B)$ and $\vdash_L x \neq x \wedge y \neq y \supset (C \leftrightarrow \langle y : x \rangle C)$. So by (PC), $\vdash_L x \neq x \wedge y \neq y \supset ((B \supset C) \leftrightarrow (\langle y : x \rangle B \supset \langle y : x \rangle C))$, and by (S⊃), $\vdash_L x \neq x \wedge y \neq y \supset ((B \supset C) \leftrightarrow \langle y : x \rangle (B \supset C))$.

4. A is $\forall z B$. We distinguish three cases.

   a) $z \notin \{x, y\}$. Then

1.	$\vdash_L x \neq x \wedge y \neq y \supset (B \leftrightarrow \langle y : x \rangle B)$	(ind. hyp.)
2.	$\vdash_L \forall z\, x \neq x \wedge \forall z\, y \neq y \supset (\forall z B \leftrightarrow \forall z \langle y : x \rangle B)$	(1, UG, UD)
3.	$\vdash_L x \neq x \wedge y \neq y \supset (\forall z B \leftrightarrow \forall z \langle y : x \rangle B)$	(2, VQ)
4.	$\vdash_L x \neq x \wedge y \neq y \supset (\forall z B \leftrightarrow \langle y : x \rangle \forall z B)$.	(3, (S∀))

   b) $z = x$. Then A is $\forall x B$, and $\vdash_L \forall x B \leftrightarrow \langle y : x \rangle \forall x B$ by (VS). So $\vdash_L x \neq x \wedge y \neq y \supset (\forall x B \leftrightarrow \langle y : x \rangle \forall x B)$ by (PC).

   c) $z = y \neq x$. Then A is $\forall y B$. Let v be a variable not in $\mathrm{Var}(A), x, y$.

1.	$\vdash_L x \neq x \wedge v \neq v \supset (B \leftrightarrow \langle v : x \rangle B)$.	(ind. hyp.)
2.	$\vdash_L \forall y\, x \neq x \wedge \forall y\, v \neq v \supset (\forall y B \leftrightarrow \forall y \langle v : x \rangle B)$.	(1, UG, UD)
3.	$\vdash_L x \neq x \wedge v \neq v \supset (\forall y B \leftrightarrow \forall y \langle v : x \rangle B)$.	(2, VQ)
4.	$\vdash_L x \neq x \wedge v \neq v \supset (\forall y B \leftrightarrow \langle v : x \rangle \forall y B)$.	(3, (S∀))
5.	$\vdash_L \langle y : v \rangle x \neq x \wedge \langle y : v \rangle v \neq v \supset (\langle y : v \rangle \forall y B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \forall y B)$.	(4, (Subˢ), (S⊃))
6.	$\vdash_L x \neq x \wedge y \neq y \supset (\langle y : v \rangle \forall y B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \forall y B)$.	(5, (VS), (SAt))
7.	$\vdash_L x \neq x \wedge y \neq y \supset (\forall y B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \forall y B)$.	(6, (VS))
8.	$\vdash_L x \neq x \wedge y \neq y \supset (\forall y B \leftrightarrow \langle y : x \rangle \forall y B)$.	(7, (SE2))

5. A is $\langle y_2 : z \rangle B$. We have four cases.

a) $z \notin \{x, y\}$ and $y_2 \neq x$. Then

1. $\vdash_L x \neq x \wedge y \neq y \supset (B \leftrightarrow \langle y : x \rangle B)$ (ind. hyp.)
2. $\vdash_L \langle y_2 : z \rangle x \neq x \wedge \langle y_2 : z \rangle y \neq y \supset (\langle y_2 : z \rangle B \leftrightarrow \langle y_2 : z \rangle \langle y : x \rangle B)$ (1, (Sub^s), (S $\supset$))
3. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y_2 : z \rangle B \leftrightarrow \langle y_2 : z \rangle \langle y : x \rangle B)$ (2, (VS))
4. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y_2 : z \rangle B \leftrightarrow \langle y : x \rangle \langle y_2 : z \rangle B)$.   (3, (SS1))

b) $z \neq x$ and $y_2 = x$. Then A is $\langle x : z \rangle B$.

1. $\vdash_L x \neq x \wedge z \neq z \supset (B \leftrightarrow \langle x : z \rangle B)$ (ind. hyp.)
2. $\vdash_L \langle y : z \rangle x \neq x \wedge \langle y : z \rangle z \neq z \supset (\langle y : z \rangle B \leftrightarrow \langle y : z \rangle \langle x : z \rangle B)$ (1, (Sub^s), (S $\supset$))
3. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y : z \rangle B \leftrightarrow \langle y : z \rangle \langle x : z \rangle B)$ (2, (SAt), $z \neq x$)
4. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y : z \rangle B \leftrightarrow \langle x : z \rangle B)$ (3, (VS), $z \neq x$)
5. $\vdash_L x \neq x \wedge y \neq y \supset (B \leftrightarrow \langle y : x \rangle B)$ (ind. hyp.)
6. $\vdash_L \langle y : z \rangle x \neq x \wedge \langle y : z \rangle y \neq y \supset (\langle y : z \rangle B \leftrightarrow \langle y : z \rangle \langle y : x \rangle B)$ (5, (Sub^s), (S $\supset$))
7. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y : z \rangle B \leftrightarrow \langle y : z \rangle \langle y : x \rangle B)$ (6, (SAt), $z \neq x$)
8. $\vdash_L x \neq x \wedge y \neq y \supset (\langle x : z \rangle B \leftrightarrow \langle y : z \rangle \langle y : x \rangle B)$ (4, 7)
9. $\vdash_L x \neq x \wedge y \neq y \supset (\langle x : z \rangle B \leftrightarrow \langle y : x \rangle \langle x : z \rangle B)$.   (8, (SS2))

c) $z = x$. Then A is $\langle y_2 : x \rangle B$, and $\vdash_L \langle y_2 : x \rangle B \leftrightarrow \langle y : x \rangle \langle y_2 : x \rangle B$ by (VS). So $\vdash_L x \neq x \wedge y \neq y \supset (\langle y_2 : x \rangle B \leftrightarrow \langle y : x \rangle \langle y_2 : x \rangle B)$ by (PC).

d) $z = y \neq x$ and $y_2 \neq x$. Then A is $\langle y_2 : y \rangle B$. Let v be a variable not in $\text{Var}(A), x, y, y_2$.

1. $\vdash_L x \neq x \wedge v \neq v \supset (B \leftrightarrow \langle v : x \rangle B)$.   (ind. hyp.)
2. $\vdash_L \langle y_2 : y \rangle x \neq x \wedge \langle y_2 : y \rangle v \neq v \supset (\langle y_2 : y \rangle B \leftrightarrow \langle y_2 : y \rangle \langle v : x \rangle B)$.   (1, (Sub^s), (S $\supset$))
3. $\vdash_L x \neq x \wedge v \neq v \supset (\langle y_2 : y \rangle B \leftrightarrow \langle y_2 : y \rangle \langle v : x \rangle B)$.   (2, (VS))
4. $\vdash_L x \neq x \wedge v \neq v \supset (\langle y_2 : y \rangle B \leftrightarrow \langle v : x \rangle \langle y_2 : y \rangle B)$.   (3, (SS1), $y_2 \neq x$)
5. $\vdash_L \langle y : v \rangle x \neq x \wedge \langle y : v \rangle v \neq v \supset (\langle y : v \rangle \langle y_2 : y \rangle B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \langle y_2 : y \rangle B)$.   (4, (Sub^s), (S $\supset$))
6. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y : v \rangle \langle y_2 : y \rangle B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \langle y_2 : y \rangle B)$.   (5, (VS), (SAt))
7. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y_2 : y \rangle B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \langle y_2 : y \rangle B)$.   (6, (VS))
8. $\vdash_L x \neq x \wedge y \neq y \supset (\langle y_2 : y \rangle B \leftrightarrow \langle y : x \rangle \langle y_2 : y \rangle B)$.   (7, (SE2))

6. *A* is $\Box B$. Let *v* be a variable not in Var(*B*).

1.	$\vdash_L x \neq x \land v \neq v \supset (B \leftrightarrow \langle v : x \rangle B)$.	(ind. hyp.)
2.	$\vdash_L \Box(x \neq x \land v \neq v) \supset (\Box B \leftrightarrow \Box \langle v : x \rangle B)$.	(1, (Nec), (K))
3.	$\vdash_L x \neq x \land v \neq v \supset \Box(x \neq x \land v \neq v)$	((NA), (EI), (Nec), (K))
4.	$\vdash_L x \neq x \land v \neq v \supset (\Box B \leftrightarrow \Box \langle v : x \rangle B)$.	(2, 3)
5.	$\vdash_L x \neq x \land v \neq v \supset (\Box B \leftrightarrow \langle v : x \rangle \Box B)$.	(4, (S$\Box$), (S$\Diamond$), $v \notin$ Var(*B*))
6.	$\vdash_L \langle y : v \rangle x \neq x \land \langle y : v \rangle v \neq v \supset (\langle y : v \rangle \Box B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \Box B)$.	(5, (Sub^s), (S$\supset$))
7.	$\vdash_L x \neq x \land y \neq y \supset (\langle y : v \rangle \Box B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \Box B)$.	(6, (SAt))
8.	$\vdash_L x \neq x \land y \neq y \supset (\Box B \leftrightarrow \langle y : v \rangle \langle v : x \rangle \Box B)$.	(7, (VS))
9.	$\vdash_L x \neq x \land y \neq y \supset (\Box B \leftrightarrow \langle y : x \rangle \Box B)$.	(8, (SE2))

Now we can prove (SC1) and (SC2). I will also show that $\langle y : x \rangle A$ and $[y/x]A$ are provably equivalent conditional on $y \neq y$.

Lemma 8.9 (Substitution conversion).
For any $\mathfrak{L}$-formula A and variables x, y,

(SC1) $\vdash_L \langle y : x \rangle A \leftrightarrow [y/x]A$, *provided y and x are modally separated in A.*

(SC2) $\vdash_L \langle y : x \rangle A \supset [y/x]A$, *provided y is really free for x in A.*

(SCN) $\vdash_L y \neq y \supset (\langle y : x \rangle A \leftrightarrow [y/x]A)$.

Proof. If *x* and *y* are the same variable, then by (SE1), $\vdash_L \langle x : x \rangle A \leftrightarrow [x/x]A$. Assume then that *x* and *y* are different variables. We first prove (SC1) and (SC2), by induction on *A*. Observe that if *A* is not a box formula $\Box B$, then by definition 7.1, *y* is really free for *x* in *A* iff *y* and *x* are modally separated in *A*, in which case *y* and *x* are also modally separated in any subformula of *A*.

1. *A* is atomic. By (SAt), $\vdash_L \langle y : x \rangle A \leftrightarrow [y/x]A$ holds without any restrictions.

2. *A* is $\neg B$. If y and x are modally separated in *A*, then by induction hypothesis, $\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$. So by (PC), $\vdash_L \neg\langle y : x \rangle B \leftrightarrow \neg[y/x]B$. By (S$\neg$) and definition 8.4, it follows that $\vdash_L \langle y : x \rangle\neg B \leftrightarrow [y/x]\neg B$.

3. *A* is $B \supset C$. If y and x are modally separated in *A*, then by induction hypothesis, $\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$ and $\vdash_L \langle y : x \rangle C \leftrightarrow [y/x]C$. By (S$\supset$), $\vdash_L \langle y : x \rangle(B \supset C) \leftrightarrow (\langle y : x \rangle B \supset \langle y : x \rangle C)$. So $\vdash_L \langle y : x \rangle(B \supset C) \leftrightarrow ([y/x]B \supset [y/x]C)$, and so $\vdash_L \langle y : x \rangle(B \supset C) \leftrightarrow [y/x](B \supset C)$ by definition 8.4.

4. *A* is $\forall z B$. We have to distinguish four cases, assuming each time that y and x are modally separated in *A*.

   a) $z \notin \{x, y\}$. By induction hypothesis, $\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$. So by (UG) and (UD), $\vdash_L \forall z \langle y : x \rangle B \leftrightarrow \forall z[y/x]B$. Since $z \notin \{x, y\}$, $\vdash_L \langle y : x \rangle\forall z B \leftrightarrow \forall z \langle y : x \rangle B$ by (S$\forall$), and $\forall z[y/x]B$ is $[y/x]\forall z B$ by definition 8.4; so $\vdash_L \langle y : x \rangle\forall z B \leftrightarrow [y/x]\forall z B$.

   b) $z = y$ and $x \notin FV(B)$. By definition 8.4, then $[y/x]\forall z B$ is $\forall y[y/x]B$.

1.	$\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$.	(induction hypothesis)
2.	$\vdash_L \forall y \langle y : x \rangle B \leftrightarrow \forall y[y/x]B$.	(1, (UG), (UD))
3.	$\vdash_L B \leftrightarrow \langle y : x \rangle B$.	((VS), $x \notin FV(B)$)
4.	$\vdash_L \forall y B \leftrightarrow \forall y \langle y : x \rangle B$.	(3, (UG), (UD))
5.	$\vdash_L \forall y B \leftrightarrow \langle y : x \rangle\forall y B$.	((VS), $x \notin FV(B)$)
6.	$\vdash_L \langle y : x \rangle\forall y B \leftrightarrow \forall y[y/x]B$.	(2, 4, 5)

   c) $z = x$ and $y \notin FV(B)$. By definition 8.4, then $[y/x]\forall z B$ is $\forall y[y/x]B$.

1.	$\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$.	(induction hypothesis)
2.	$\vdash_L \forall y \langle y : x \rangle B \leftrightarrow \forall y[y/x]B$.	(1, (UG), (UD))
3.	$\vdash_L \forall x B \leftrightarrow \forall y \langle y : x \rangle B$.	((SBV), $y \notin FV(B)$)
4.	$\vdash_L \forall x B \leftrightarrow \langle y : x \rangle\forall x B$.	(VS)
5.	$\vdash_L \langle y : x \rangle\forall x B \leftrightarrow \forall y[y/x]B$.	(2, 3, 4)

116

d) $z = x$ and $y \in \mathrm{FV}(B)$, or $z = y$ and $x \in \mathrm{FV}(B)$. By definition 8.4, then $[y/x]\forall zB$ is $\forall v[y/x][v/z]B$ for some variable $v \notin \mathrm{Var}(B) \cup \{x, y\}$. Since v and z are modally separated in B, by induction hypothesis $\vdash_L \langle v : z \rangle B \leftrightarrow [v/z]B$. So by (UG) and (UD), $\vdash_L \forall v \langle v : z \rangle B \leftrightarrow \forall v[v/z]B$. By (SBV), $\vdash_L \forall zB \leftrightarrow \forall v \langle v : z \rangle B$. So $\vdash_L \forall zB \leftrightarrow \forall v[v/z]B$. Moreover, as $z \in \{x, y\}$, y and x are modally separated in $[v/z]B$. So by induction hypothesis, $\vdash_L \langle y : x \rangle [v/z]B \leftrightarrow [y/x][v/z]B$. Then

1. $\vdash_L \forall zB \leftrightarrow \forall v[v/z]B$ (as just shown)
2. $\vdash_L \langle y : x \rangle \forall zB \leftrightarrow \langle y : x \rangle \forall v[v/z]B$ (1, (Sub^s), (S$\neg$), (S$\supset$))
3. $\vdash_L \langle y : x \rangle \forall v[v/z]B \leftrightarrow \forall v \langle y : x \rangle [v/z]B$.   (S$\forall$)
4. $\vdash_L \langle y : x \rangle \forall zB \leftrightarrow \forall v \langle y : x \rangle [v/z]B$.   (2, 3)
5. $\vdash_L \langle y : x \rangle [v/z]B \leftrightarrow [y/x][v/z]B$.   (induction hypothesis)
6. $\vdash_L \forall v \langle y : x \rangle [v/z]B \leftrightarrow \forall v[y/x][v/z]B$.   (5, (UG), (UD))
7. $\vdash_L \langle y : x \rangle \forall zB \leftrightarrow \forall v[y/x][v/z]B$.   (4, 6)

5. A is $\langle y_2 : z \rangle B$. Again we have four cases, assuming x and y are modally separated in A.

   a) $z \notin \{x, y\}$. By definition 8.4, then $[y/x]\langle y_2 : z \rangle B$ is $\langle [y/x]y_2 : z \rangle [y/x]B$.

   1. $\vdash \langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle [y/x]y_2 : z \rangle \langle y : x \rangle B$ ((SS1) or (SS2))
   2. $\vdash \langle y : x \rangle B \leftrightarrow [y/x]B$ (induction hypothesis)
   3. $\vdash \langle [y/x]y_2 : z \rangle (\langle y : x \rangle B \leftrightarrow [y/x]B)$ (2, (Sub^s))
   4. $\vdash \langle [y/x]y_2 : z \rangle \langle y : x \rangle B \leftrightarrow \langle [y/x]y_2 : z \rangle [y/x]B$ (3, (S$\supset$), (S$\neg$))
   5. $\vdash \langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle [y/x]y_2 : z \rangle [y/x]B$.   (1, 4)

   b) $z = y$ and $x \notin \mathrm{FV}(B)$. By definition 8.4, then $[y/x]\langle y_2 : z \rangle B$ is $\langle [y/x]y_2 : y \rangle [y/x]B$. By induction hypothesis, $\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$. So by (Sub^s) and (S$\supset$), $\vdash_L \langle [y/x]y_2 : y \rangle \langle y : x \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle [y/x]B$. If $y_2 = x$, then

$\vdash_L \langle y : x \rangle\langle y_2 : y \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle\langle y : x \rangle B$ by (SS2). If $y_2 \neq x$, then

1. $\vdash_L \langle y_2 : y \rangle B \leftrightarrow \langle y : x \rangle\langle y_2 : y \rangle B$ $((\text{VS}), x \notin \text{FV}(\langle y_2 : y \rangle B))$
2. $\vdash_L B \leftrightarrow \langle y : x \rangle B$ $((\text{VS}), x \notin \text{FV}(B))$
3. $\vdash_L \langle y_2 : y \rangle B \leftrightarrow \langle y_2 : y \rangle\langle y : x \rangle B$ $(1, (\text{Sub}^s), (\text{S}\supset))$
4. $\vdash_L \langle y : x \rangle\langle y_2 : y \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle\langle y : x \rangle B$ $(1, 3)$

So either way $\vdash_L \langle y : x \rangle\langle y_2 : y \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle\langle y : x \rangle B$. So $\vdash_L \langle y : x \rangle\langle y_2 : y \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle[y/x]B$.

c) $z = x$ and $y \notin \text{FV}(B)$. By definition 8.4, then $[y/x]\langle y_2 : z \rangle B$ is $([y/x]y_2 : y)[y/x]B$. By induction hypothesis, $\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$. So by (Sub^s) and $(\text{S}\supset)$, $\vdash_L \langle [y/x]y_2 : y \rangle\langle y : x \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle[y/x]B$. Since $y \notin \text{FV}(B)$, by (SE2), $\vdash_L \langle [y/x]y_2 : y \rangle\langle y : x \rangle B \leftrightarrow \langle [y/x]y_2 : x \rangle B$. Moreover, $\vdash_L \langle [y/x]y_2 : x \rangle B \leftrightarrow \langle y : x \rangle\langle y_2 : x \rangle B$ by either (VS) (if $x \neq y_2$) or by (SE1), (Sub^s) and $(\text{S}\supset)$ (if $x = y_2$). So $\vdash_L \langle y : x \rangle\langle y_2 : x \rangle B \leftrightarrow \langle [y/x]y_2 : y \rangle\langle y : x \rangle B$.

d) $z = x$ and $y \in \text{FV}(B)$, or $z = y$ and $x \in \text{FV}(B)$. By definition 8.4, then $[y/x]\langle y_2 : z \rangle B$ is $\langle [y/x]y_2 : v \rangle[y/x][v/z]B$, where $v \notin \text{Var}(B) \cup \{x, y, y_2\}$.

1. $\vdash \langle v : z \rangle B \leftrightarrow [v/z]B$ (induction hypothesis)
2. $\vdash \langle y_2 : v \rangle\langle v : z \rangle B \leftrightarrow \langle y_2 : v \rangle[v/z]B$ $(1, (\text{Sub}^s), (\text{S}\supset), (\text{S}\neg))$
3. $\vdash \langle y_2 : z \rangle B \leftrightarrow \langle y_2 : v \rangle\langle v : z \rangle B$ (SE2)
4. $\vdash \langle y_2 : z \rangle B \leftrightarrow \langle y_2 : v \rangle[v/z]B$ $(2, 3)$

Since $z \in \{x, y\}$, x and y are modally separated in $[v/z]B$. So:

5.	$\vdash \langle y : x \rangle [v/z]B \leftrightarrow [y/x][v/z]B$	(ind. hyp.)
6.	$\vdash \langle [y/x]y_2 : v \rangle \langle y : x \rangle [v/z]B \leftrightarrow \langle [y/x]y_2 : v \rangle [y/x][v/z]B$	(5, (Sub^s), (S⊃))
7.	$\vdash \langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle y : x \rangle \langle y_2 : v \rangle [v/z]B$	(4, (Sub^s), (S⊃))
8.	$\vdash \langle y : x \rangle \langle y_2 : v \rangle [v/z]B \leftrightarrow \langle [y/x]y_2 : v \rangle \langle y : x \rangle [v/z]B$	((SS1) or (SS2))
9.	$\vdash \langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle [y/x]y_2 : v \rangle \langle y : x \rangle [v/z]B$	(7, 8)
10.	$\vdash \langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle [y/x]y_2 : v \rangle [y/x][v/z]B$	(6, 9)

6. *A* is $\Box B$. For (SC1), assume x and y are modally separated in *A*. Then they are also modally separated in *B*, so by induction hypothesis, $\vdash_L \langle y : x \rangle B \leftrightarrow [y/x]B$. By (Nec) and (K), then $\vdash_L \Box \langle y : x \rangle B \leftrightarrow \Box[y/x]B$. By (S$\Box$), $\vdash_L \langle y : x \rangle \Box B \supset \Box \langle y : x \rangle B$. Since at most one of x, y is free in *B*, by (S$\Diamond$), $\vdash_L \langle y : x \rangle \Diamond \neg B \supset \Diamond \langle y : x \rangle \neg B$; so $\vdash_L \Box \langle y : x \rangle B \supset \langle y : x \rangle \Box B$ (by (S¬), (Sub^s), (S⊃), (Nec), (K)). So $\vdash_L \langle y : x \rangle \Box B \leftrightarrow \Box[y/x]B$. Since $\Box[y/x]B$ is $[y/x]\Box B$ by definition 8.4, this means that $\vdash_L \langle y : x \rangle \Box B \leftrightarrow [y/x]\Box B$.

   For (SC2), assume y is really free for x in $\Box B$. Then y is really free for x in *B*, so by induction hypothesis, $\vdash \langle y : x \rangle B \supset [y/x]B$. By (Nec) and (K), then $\vdash \Box \langle y : x \rangle B \supset \Box[y/x]B$. By (S$\Box$), $\vdash \langle y : x \rangle \Box B \supset \Box \langle y : x \rangle B$. So $\vdash \langle y : x \rangle \Box B \supset \Box[y/x]B$.

Here is the proof for (SCN). The first three clauses are very similar.

1. *A* is atomic. Then $\vdash_L \langle y : x \rangle A \leftrightarrow [y/x]A$ as we've seen above, and so $\vdash_L y \neq y \supset (\langle y : x \rangle A \leftrightarrow [y/x]A)$ by (PC).

2. *A* is $\neg B$. By induction hypothesis, $\vdash_L y \neq y \supset (\langle y : x \rangle B \leftrightarrow [y/x]B)$. So by (PC), $\vdash_L y \neq y \supset (\neg \langle y : x \rangle B \leftrightarrow \neg [y/x]B)$. By (S¬) and definition 8.4, it follows that $\vdash_L y \neq y \supset (\langle y : x \rangle \neg B \leftrightarrow [y/x]\neg B)$.

3. *A* is $B \supset C$. By induction hypothesis, $\vdash_L y \neq y \supset (\langle y : x \rangle B \leftrightarrow [y/x]B)$ and $\vdash_L y \neq y \supset (\langle y : x \rangle C \leftrightarrow [y/x]C)$. By (S⊃), $\vdash_L y \neq y \supset (\langle y : x \rangle (B \supset C) \leftrightarrow (\langle y : x \rangle B \supset \langle y : x \rangle C))$. So $\vdash_L y \neq y \supset (\langle y : x \rangle (B \supset C) \leftrightarrow ([y/x]B \supset [y/x]C))$, and so $\vdash_L y \neq y \supset (\langle y : x \rangle (B \supset C) \leftrightarrow [y/x](B \supset C))$ by definition 8.4.

4. *A* is $\forall zB$. If $z \notin \{x, y\}$, then by induction hypothesis, $\vdash_L y \neq y \supset (\langle y : x \rangle B \leftrightarrow [y/x]B)$. So by (UG) and (UD), $\vdash_L \forall z\, y \neq y \supset (\forall z \langle y : x \rangle B \leftrightarrow \forall z[y/x]B)$. Since $z \notin \{x, y\}$, $\vdash_L \langle y : x \rangle \forall zB \leftrightarrow \forall z \langle y : x \rangle B$ by (S$\forall$), and $\vdash_L y \neq y \supset \forall z\, y \neq y$ by (VQ), and $\forall z[y/x]B$ is $[y/x]\forall zB$ by definition 8.4; so $\vdash_L y \neq y \supset (\langle y : x \rangle \forall zB \leftrightarrow [y/x]\forall zB)$.

   Alternatively, if $z \in \{x, y\}$, then either x or y is not free in A, and thus x and y are modally separated in A. By (SC2), then $\vdash_L \langle y : x \rangle \forall zB \leftrightarrow [y/x]\forall zB$, and so by (PC), $\vdash_L y \neq y \supset (\langle y : x \rangle \forall zB \leftrightarrow [y/x]\forall zB)$.

5. *A* is $\langle y_2 : z \rangle B$. If $z \notin \{x, y\}$, then by induction hypothesis, $\vdash_L y \neq y \supset (\langle y : x \rangle B \leftrightarrow [y/x]B)$. So by (Sub^s) and (S$\supset$), $\vdash_L \langle [y/x]y_2 : z \rangle y \neq y \supset (\langle [y/x]y_2 : z \rangle \langle y : x \rangle B \leftrightarrow \langle [y/x]y_2 : z \rangle [y/x]B)$. By (VS), $\langle [y/x]y_2 : z \rangle y \neq y \leftrightarrow y \neq y$. And by (SS1) or (SS2), $\langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle [y/x]y_2 : z \rangle \langle y : x \rangle B$. So $\vdash_L y \neq y \supset (\langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow \langle [y/x]y_2 : z \rangle [y/x]B)$. But by definition 8.4, $[y/x]\langle y_2 : z \rangle B$ is $\langle [y/x]y_2 : y \rangle [y/x]B$.

   Alternatively, if $z \in \{x, y\}$, then either x or y is not free in A, and thus x and y are modally separated in A. By (SC2), then $\vdash_L \langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow [y/x]\langle y_2 : z \rangle B$, and so by (PC), $\vdash_L y \neq y \supset (\langle y : x \rangle \langle y_2 : z \rangle B \leftrightarrow [y/x]\langle y_2 : z \rangle B)$.

6. *A* is $\Box B$. Then

1.	$\vdash_L y \neq y \supset (\langle y : x \rangle B \leftrightarrow [y/x]B)$.	(ind. hyp.)
2.	$\vdash_L \Box y \neq y \supset (\Box \langle y : x \rangle B \leftrightarrow \Box [y/x]B)$.	(1, (Nec), (K))
3.	$\vdash_L y \neq y \supset \Box y \neq y$.	((=R) or (NA), (EI) and (Nec))
4.	$\vdash_L y \neq y \supset (\Box \langle y : x \rangle B \leftrightarrow \Box [y/x]B)$.	(2, 3)
5.	$\vdash_L y \neq y \supset \langle y : x \rangle (x \neq x \wedge y \neq y)$	((SAt), (S$\supset$), (S$\neg$))
6.	$\vdash_L (x \neq x \wedge y \neq y) \supset \Box(x \neq x \wedge y \neq y)$.	((=R) or (NA), (EI), (Nec) and (K))
7.	$\vdash_L \Box(x \neq x \wedge y \neq y) \supset (\Box B \leftrightarrow \Box \langle y : x \rangle B)$.	((SEV), (Nec), (K))
8.	$\vdash_L (x \neq x \wedge y \neq y) \supset (\Box B \leftrightarrow \Box \langle y : x \rangle B)$.	(6, 7)
9.	$\vdash_L \langle y : x \rangle (x \neq x \wedge y \neq y) \supset (\langle y : x \rangle \Box B \leftrightarrow \langle y : x \rangle \Box \langle y : x \rangle B)$.	(8, (Sub^s), (S$\supset$))
10.	$\vdash_L \langle y : x \rangle (x \neq x \wedge y \neq y) \supset (\langle y : x \rangle \Box B \leftrightarrow \Box \langle y : x \rangle B)$.	(9, (VS))
11.	$\vdash_L y \neq y \supset (\langle y : x \rangle \Box B \leftrightarrow \Box \langle y : x \rangle B)$.	(7, 10)
12.	$\vdash_L y \neq y \supset (\langle y : x \rangle \Box B \leftrightarrow [y/x]\Box B)$.	(4, 13, def. 8.4)

For the next lemma, we extend the definition of alphabetic variants (definition 3.6) to formulas with the substitution operator.

Definition 8.7 (Alphabetic variant).
A formula A' is an *alphabetic variant of* a formula A if one of the following conditions is satisfied.

1. $A = A'$.

2. $A = \neg B$, $A' = \neg B'$, and B' is an alphabetic variant of B.

3. $A = B \supset C$, $A' = B' \supset C'$, and B', C' are alphabetic variants of B, C, respectively.

4. $A = \forall x B$, $A' = \forall z[z/x]B'$, B' is an alphabetic variant of B, and either $z = x$ or $z \notin \text{Var}(B')$.

5. $A = \langle y : x \rangle B$, $A' = \langle y : z \rangle[z/x]B'$, B' is an alphabetic variant of B, and either $z = x$ or $z \notin \text{Var}(B')$.

6. $A = \Box B$, $A' = \Box B'$, and B' is an alphabetic variant of A'.

Lemma 8.10 (Syntactic alpha-conversion).
If A, A' are $\mathfrak{L}$-formulas, and A' is an alphabetic variant of A, then

(AC) $\vdash_L A \leftrightarrow A'$.

Proof. by induction on A.

1. A is atomic. Then $A = A'$ and $\vdash_L A \leftrightarrow A'$ by (Taut).

2. A is $\neg B$. Then A' is $\neg B'$ with B' an alphabetic variant of B. By induction hypothesis, $\vdash_L B \leftrightarrow B'$. By (PC), $\vdash_L \neg B \leftrightarrow \neg B'$.

3. A is $B \supset C$. Then A' is $B' \supset C'$ with B', C' alphabetic variants of B, C, respectively. By induction hypothesis, $\vdash_L B \leftrightarrow B'$ and $\vdash_{sC} C \leftrightarrow C'$. By (PC), then $\vdash_L (B \supset C) \leftrightarrow (B' \supset C')$.

4. A is $\forall x B$. Then A' is either $\forall x B'$ or $\forall z [z/x] B'$, where B' is an alphabetic variant of B and $z \notin \mathrm{Var}(B')$. Assume first that A' is $\forall x B'$. By induction hypothesis, $\vdash_L B \leftrightarrow B'$. So by (UG) and (UD), $\vdash_L \forall x B \leftrightarrow \forall x B'$.

   Alternatively, assume A' is $\forall z [z/x] B'$ and $z \notin \mathrm{Var}(B')$. Since B' differs from B at most in renaming bound variables, if z were free in B, then $z \in \mathrm{Var}(B')$. So z is not free in B. Then

1.	$\vdash_L B \leftrightarrow B'.$	induction hypothesis
2.	$\vdash_L \langle z : x \rangle B \leftrightarrow \langle z : x \rangle B'.$	$(1, (\text{Sub}^s), (\text{S}\neg))$
3.	$\vdash_L \langle z : x \rangle B' \leftrightarrow [z/x] B'.$	$((\text{SC1}), z \notin \mathrm{Var}(B'))$
4.	$\vdash_L \langle z : x \rangle B \leftrightarrow [z/x] B'.$	$(2, 3)$
5.	$\vdash_L \forall z \langle z : x \rangle B \leftrightarrow \forall z [z/x] B'.$	$(4, (\text{UG}), (\text{UD}))$
6.	$\vdash_L \forall x B \leftrightarrow \forall z \langle z : x \rangle B.$	$((\text{SBV}), z \text{ not free in } B)$
7.	$\vdash_L \forall x B \leftrightarrow \forall z [z/x] B'.$	$(5, 6)$

5. A is $\langle y : x \rangle B$. Then A' is either $\langle y : x \rangle B'$ or $\langle y : z \rangle [z/x] B'$, where B' is an alphabetic variant of B and $z \notin \mathrm{Var}(B)$. Assume first that A' is $\langle y : x \rangle B'$. By induction hypothesis, $\vdash_L B \leftrightarrow B'$. So by (Sub^s) and (S $\supset$), $\vdash_L \langle y : x \rangle B \leftrightarrow \langle y : x \rangle B'$.

   Alternatively, assume A' is $\langle y : z \rangle [z/x] B'$ and $z \notin \mathrm{Var}(B')$. Again, it

follows that z is not free in B. So

1. $\vdash_L B \leftrightarrow B'$.   induction hypothesis

2. $\vdash_L \langle z : x \rangle B \leftrightarrow \langle z : x \rangle B'$.   (1, (Sub^s), (S $\supset$))

3. $\vdash_L \langle z : x \rangle B' \leftrightarrow [z/x]B'$.   ((SC1), $z \notin \mathrm{Var}(B')$)

4. $\vdash_L \langle z : x \rangle B \leftrightarrow [z/x]B'$.   (2, 3)

5. $\vdash_L \langle y : z \rangle \langle z : x \rangle B \leftrightarrow \langle y : z \rangle [z/x]B'$.   (4, (Sub^s), (S $\supset$))

6. $\vdash_L \langle y : z \rangle \langle z : x \rangle B \leftrightarrow \langle y : x \rangle B$.   ((SE2), z not free in B)

7. $\vdash_L \langle y : x \rangle B \leftrightarrow \langle y : z \rangle [z/x]B'$.   (5, 6)

6. *A* is $\Box A'$. Then *B* is $\Box B'$ with B' an alphabetic variant of A'. By induction hypothesis, $\vdash_L A' \leftrightarrow B'$. Then by (Nec), $\vdash_L \Box(A' \leftrightarrow B')$, and by (K), $\vdash_L \Box A' \leftrightarrow \Box B'$.

Theorem 8.11 (Substitution and non-substitution logics).
For any $\mathfrak{L}$-formula A and variables x, y,

(FUI*) $\vdash_L \forall x A \supset (Ey \supset [y/x]A)$, *provided y is really free for x in A*,

(LL*) $\vdash_L x = y \supset A \supset [y/x]A$, *provided y is really free for x in A*,

(Sub*)   *if $\vdash_L A$, then $\vdash_L [y/x]A$, provided y is really free for x in A*.

It follows that FK* $\subseteq$ FK^s *and* NK* $\subseteq$ NK^s.

Proof. Assume *y* is really free for *x* in *A*. Then by (SC2), $\vdash_L \langle y : x \rangle A \supset [y/x]A$. By (FUI^s), $\vdash_L \forall x A \supset (Ey \supset \langle y : x \rangle A)$, so by (PC), $\vdash_L \forall x A \supset (Ey \supset [y/x]A)$. Similarly, by (LL^s), $\vdash_L x = y \supset A \supset \langle y : x \rangle A$, so by (PC), $\vdash_L x = y \supset A \supset [y/x]A$. Finally, by (Sub^s), if $\vdash_L A$, then $\vdash_L \langle y : x \rangle A$, so then $\vdash_L [y/x]A$ by (PC).

Lemma 8.12 (Symmetry and transitivity of identity).
For any $\mathfrak{L}$-variables x, y, z,

$(=S)$ $\vdash_L x = y \supset y = x;$

$(=T)$ $\vdash_L x = y \supset y = z \supset x = z.$

Proof. Immediate from theorem 8.11 and lemma 7.12.

Lemma 8.13 (Variations on Leibniz' Law).
If A is an $\mathfrak{L}$-formula and x, y, y' are $\mathfrak{L}$-variables, then

(LV1) $\vdash_L x = y \supset \langle y : x \rangle A \supset A.$

(LV2) $\vdash_L y = y' \supset \langle y : x \rangle A \supset [y'/x]A$, *provided y' is really free for x in A.*

Proof. (LV1). Let z be an $\mathfrak{L}$-variable not in $\mathrm{Var}(A)$. Then

1.	$\vdash_L x = z \supset \langle z : x \rangle A \supset \langle x : z \rangle \langle z : x \rangle A.$	(LL^s)
2.	$\vdash_L x = z \supset \langle z : x \rangle A \supset \langle x : x \rangle A.$	$(1, (SE2), z \notin \mathrm{Var}(A))$
3.	$\vdash_L x = z \supset \langle z : x \rangle A \supset A.$	$(2, (SE1))$
4.	$\vdash_L \langle y : z \rangle x = z \supset \langle y : z \rangle \langle z : x \rangle A \supset \langle y : z \rangle A.$	$(3, (VS), (S \supset))$
5.	$\vdash_L x = z \supset \langle y : z \rangle \langle z : x \rangle A \supset \langle y : z \rangle A.$	$(4, (SAt))$
6.	$\vdash_L x = z \supset \langle y : x \rangle A \supset \langle y : z \rangle A.$	$(5, (SE2), z \notin \mathrm{Var}(A))$
7.	$\vdash_L x = z \supset \langle y : x \rangle A \supset A.$	$(6, (VS), z \notin \mathrm{Var}(A))$.

(LV2).

1. $\vdash_L x = y \wedge y = y' \supset x = y'.$ $(= T)$
2. $\vdash_L A \wedge x = y' \supset [y'/x]A.$ $((LL^*), y' \text{ m.f. in } A)$
3. $\vdash_L A \wedge x = y \wedge y = y' \supset [y'/x]A.$ $(1, 2)$
4. $\vdash_L \langle y : x \rangle A \wedge \langle y : x \rangle x = y \wedge \langle y : x \rangle y = y' \supset \langle y : x \rangle [y'/x]A.$ $(3, (\mathrm{Sub^s}), (\mathrm{S}\neg), (\mathrm{S}\supset))$
5. $\vdash_L y = y \supset \langle y : x \rangle x = y.$ (SAt)
6. $\vdash_L y = y' \supset y = y.$ $((LL^*), (= \mathrm{S}))$
7. $\vdash_L y = y' \supset \langle y : x \rangle y = y'.$ (VS)
8. $\vdash_L \langle y : x \rangle A \wedge y = y' \supset \langle y : x \rangle [y'/x]A.$ $(4, 5, 6, 7)$
9. $\vdash_L \langle y : x \rangle [y'/x]A \supset [y'/x]A.$ (VS)
10. $\vdash_L \langle y : x \rangle A \wedge y = y' \supset [y'/x]A.$ $(8, 9)$

Lemma 8.14 (Leibniz' Law with sequences).
For any $\mathfrak{L}$-formula A and variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ such that the $x_1, \ldots, x_n$ are pairwise distinct,

$$(\mathrm{LL_n}) \;\; \vdash_L x_1 = y_1 \wedge \ldots \wedge x_n = y_n \supset A \supset \langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A.$$

Proof. For $n = 1$, $(\mathrm{LL_n})$ is $(\mathrm{LL^s})$. Assume then that $n > 1$. To keep formulas in the following proof at a managable length, let $\underline{\phi(i)}$ abbreviate the sequence $\phi(1), \ldots, \phi(n-1)$. For example, $\langle \underline{y_i : x_i} \rangle$ is $\langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle$. Let z

be the alphabetically first variable not in A or $x_1, \ldots, x_n$. Now

1. $\vdash_L x_n = y_n \supset \langle \underline{y_i : x_i} \rangle A \supset \langle y_n : x_n \rangle \langle \underline{y_i : x_i} \rangle A.$ (LL^s)

2. $\vdash_L \langle y_n : x_n \rangle \langle \underline{y_i : x_i} \rangle A \supset \langle y_n : z \rangle \langle z : x_n \rangle \langle \underline{y_i : x_i} \rangle A.$ (SE1)

3. $\vdash_L \langle z : x_n \rangle \langle \underline{y_i : x_i} \rangle A \supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A.$ ((SS1) or (SS2))

4. $\vdash_L \langle y_n : z \rangle \langle z : x_n \rangle \langle \underline{y_i : x_i} \rangle A$
   $\supset \langle y_n : z \rangle \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A.$ (3, (Sub^s), (S$\supset$))

5. $\vdash_L x_n = y_n \supset \langle \underline{y_i : x_i} \rangle A \supset \langle y_n : z \rangle \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A.$ (1, 2, 4)

6. $\vdash_L x_n = z \supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A$
   $\supset \langle z : x_n \rangle \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A.$ (LL^s)

7. $\vdash_L x_n = z \supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A$
   $\supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle \langle z : x_n \rangle A.$ (6, (SS1))

8. $\vdash_L z = x_n \supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle \langle z : x_n \rangle A$
   $\supset \langle x_n : z \rangle \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle \langle z : x_n \rangle A.$ (LL^s)

9. $\vdash_L z = x_n \supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle \langle z : x_n \rangle A$
   $\supset \langle \underline{y_i : x_i} \rangle \langle x_n : z \rangle \langle z : x_n \rangle \langle z : x_n \rangle A.$ (8, (SS1), (SS2))

10. $\vdash_L \langle x_n : z \rangle \langle z : x_n \rangle \langle z : x_n \rangle A \leftrightarrow \langle z : x_n \rangle A$ ((SE1), (SE2))

11. $\vdash_L \langle \underline{y_i : x_i} \rangle \langle x_n : z \rangle \langle z : x_n \rangle \langle z : x_n \rangle A \supset \langle \underline{y_i : x_i} \rangle \langle z : x_n \rangle A$ (10, (Sub^s), (S$\supset$))

12. $\vdash_L z = x_n \supset x_n = z$ (=S)

13. $\vdash_L z = x_n \supset \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A \supset \langle \underline{y_i : x_i} \rangle \langle z : x_n \rangle A.$ (7, 9, 11, 12)

14. $\vdash_L x_n = y_n \supset \langle y_n : z \rangle z = x_n$ ((=S), (SAt))

15. $\vdash_L x_n = y_n \supset \langle y_n : z \rangle \langle \underline{[z/x_n]y_i : x_i} \rangle \langle z : x_n \rangle A$
    $\supset \langle y_n : z \rangle \langle \underline{y_i : x_i} \rangle \langle z : x_n \rangle A.$ 13, 14, (Sub^s), (S$\supset$)

16. $\vdash_L x_n = y_n \supset \langle \underline{y_i : x_i} \rangle A \supset \langle y_n : z \rangle \langle \underline{y_i : x_i} \rangle \langle z : x_n \rangle A.$ 5, 15

17. $\vdash_L x_1 = y_1 \wedge \ldots \wedge x_{n-1} = y_{n-1} \supset A \supset \langle \underline{y_i : x_i} \rangle A.$ (induction hypothesis)

18. $\vdash_L x_1 = y_1 \wedge \ldots \wedge x_n = y_n \supset A \supset \langle y_n : z \rangle \langle \underline{y_i : x_i} \rangle \langle z : x_n \rangle A.$ (16, 17)

19. $\vdash_L x_1 = y_1 \wedge \ldots \wedge x_n = y_n \supset A \supset \langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A.$ (18, def. 8.3)

Lemma 8.15 (Closure under injective substitutions).
For any $\mathfrak{L}$-formula A and injective substitution τ on $\mathfrak{L}$,

$(\mathrm{Sub^t})\ \vdash_L A$ *iff* $\vdash_L A^\tau$.

Proof. The proof is almost exactly as in lemma 7.13. We only need to add a clause to the proof that that $A^\tau = [x_n^\tau/v_n] \dots [x_2^\tau/v_2][x_1^\tau/x_1][v_2/x_2] \dots [v_n/x_n]A$, where $v_2, \dots, v_n$ are distinct variables not in A or A^τ. The proof is by induction on the subformulas B of A, ordered by complexity. Recall that Σ abbreviates $[x_n^\tau/v_n] \dots [x_2^\tau/v_2][x_1^\tau/x_1][v_2/x_2] \dots [v_n/x_n]$.

Now assume B is $\langle y : z\rangle C$. By induction hypothesis, $\Sigma C = C^\tau$. Since τ is injective, it follows by definition 8.4 that $\Sigma\langle y : z\rangle C$ is $\langle \Sigma y : \Sigma z\rangle\Sigma C$, and $(\langle y : z\rangle C)^\tau$ is $\langle y^\tau : z^\tau\rangle C^\tau$. It is easy to verify that $\Sigma y = y^\tau$ and $\Sigma z = z^\tau$.

9 Canonical models for non-functional logics

9.1 Preview

Establishing completeness for the functional base logics from chapter 3 was easy. For nonfunctional logics, the situation is more complicated. Here's why.

As before, we assume that singular terms denote equivalence classes of terms. Any counterpart relation in a canonical model will therefore be a relation between variable classes. We would like to represent such relations by substitutions, as in the functional case. In non-functional models, it is not obvious that this can be done. Suppose $\{x\}$ at w has two C-counterparts $\{u\}, \{v\}$ at w'. Then there is no substitution that maps x to both u and v. We could try to work with substitution *relations*. But this gets messy. My strategy is to say that a class $[x]_w$ at w only has two C-counterparts at w' if it contains at least two variables: $\{x, y\}$ can have both $\{u\}$ and $\{v\}$ as C-counterparts, relative to a substitution that maps x to u and y to v.

The next problem we face is that the substitution lemma doesn't hold for non-functional logics. In chapter 4, we used the substitution lemma in the proof of the truth lemma. Knowing that $w, g_w \Vdash \Box A$ iff $w', g_{w'} \circ \sigma \Vdash A$ for all w' and σ such that $\{\sigma(X) : \Box X \in w\} \subseteq w'$, we'd like to infer that $w', g_{w'} \Vdash \sigma(A)$ for all such w' and σ so that we can apply the induction hypothesis.

If we have a substitution operator in the language, we can replace $\sigma(A)$ by $\langle y_1 \dots y_n / x_1 \dots x_n \rangle A$. Without a substitution operator, we can get around the problem by stipulating that all relevant substitutions σ must be injective. Injective substitutions – which we'll henceforth call *transformations* – make capturing impossible: for the free variable y in $\forall x A(y)$ to be captured by the initial quantifier $\forall x$ after substitution, x and y have to be replaced by the same variable. Indeed, definition 3.1 has been chosen to entail that if σ is injective then $\sigma(A)$ is simply A with all variables simultaneously replaced

by their σ-value.

> **Lemma 9.1 (Transformation Lemma).**
> *For any counterpart model $\mathfrak{M}$, world w in $\mathfrak{M}$, assignment g on U_w, formula A, and transformation τ,*
>
> $$\mathfrak{M}, w, g \circ \tau \Vdash A \text{ iff } \mathfrak{M}, w, g \Vdash \tau(A).$$

Proof. By induction on A.

1. A is $Px_1 \dots x_n$. $\mathfrak{M}, w, g \circ \tau \Vdash Px_1 \dots x_n$ iff $\langle (g \circ \tau)(x_1), \dots, (g \circ \tau)(x_n) \rangle \in I_w(P)$ by definition 2.9, iff $\langle g(x_1^\tau), \dots, g(x_n^\tau) \rangle \in I_w(P)$, iff $\mathfrak{M}, w, g \Vdash \tau(Px_1 \dots x_n)$ by definition 3.1.

2. A is $\neg B$. $\mathfrak{M}, w, g \circ \tau \Vdash \neg B$ iff $\mathfrak{M}, w, g \circ \tau \nVdash B$ by definition 2.9, iff $\mathfrak{M}, w, g \nVdash \tau(B)$ by induction hypothesis, iff $\mathfrak{M}, w, g \Vdash \neg\tau(B)$ by definition 2.9, iff $\mathfrak{M}, w, g \Vdash \tau(\neg B)$ by definition 3.1.

3. A is $B \supset C$. Analogous.

4. A is $\langle y : x \rangle B$. By definition 8.2, $\mathfrak{M}, w, g \circ \tau \Vdash \langle y : x \rangle B$ iff $w, (g \circ \tau)^{[y/x]} \Vdash B$. Now $(g \circ \tau)^{[y/x]}(x) = (g \circ \tau)(y) = g^{[y^\tau/x^\tau]}(x^\tau) = (g^{[y^\tau/x^\tau]} \circ \tau)(x)$. Also, for any variable $z \neq x$, $(g \circ \tau)^{[y/x]}(z) = (g \circ \tau)(z) = g^{[y^\tau/x^\tau]}(z^\tau$ (because $z^\tau \neq x^\tau$, by injectivity of $\tau) = (g^{[y^\tau/x^\tau]} \circ \tau)(z)$. So $(g \circ \tau)^{[y/x]} = g^{[y^\tau/x^\tau]} \circ \tau$. And so $\mathfrak{M}, w, (g \circ \tau)^{[y/x]} \Vdash B$ iff $\mathfrak{M}, w, g^{[y^\tau/x^\tau]} \circ \tau \Vdash B$. The latter holds iff $\mathfrak{M}, w, g^{[y^\tau/x^\tau]} \Vdash \tau(B)$ by induction hypothesis, iff $\mathfrak{M}, w, g \Vdash \langle y^\tau : x^\tau \rangle \tau(B)$ by definition 8.2, iff $\mathfrak{M}, w, g \Vdash \tau(\langle y : x \rangle B)$ by definition 3.1.

5. A is $\forall x B$. Assume $\mathfrak{M}, w, g \circ \tau \nVdash \forall x B$. Then $\mathfrak{M}, w, g^{x \mapsto d} \nVdash B$ for some $d \in D_w$. We have $(g^{x^\tau \mapsto d} \circ \tau)(x) = g^{x \mapsto d}(x)$. For any variable $z \neq x$, we also have $(g^{x^\tau \mapsto d} \circ \tau)(z) = g^{x \mapsto d}(z)$ (because $z^\tau \neq x^\tau$, by injectivity of τ). So $g^{x^\tau \mapsto d} \circ \tau = g^{x \mapsto d}$ and $\mathfrak{M}, w, g^{x^\tau \mapsto d} \nVdash B$. And so $\mathfrak{M}, w, g^{x^\tau \mapsto d} \circ \tau \nVdash B$. By induction hypothesis, it follows that $\mathfrak{M}, w, g^{x^\tau \mapsto d} \nVdash \tau(B)$. By definition 2.9, this means that $\mathfrak{M}, w, g \nVdash \forall x^\tau \tau(B)$. Hence $\mathfrak{M}, w, g \nVdash \tau(\forall x B)$ by definition 3.1.

130

In the other direction, assume $\mathfrak{M}, w, g \not\Vdash \tau(\forall x B)$, and thus $\mathfrak{M}, w, g \not\Vdash \forall x^\tau \tau(B)$. Then $\mathfrak{M}, w, g^{x^\tau \mapsto d} \not\Vdash \tau(B)$ for some $d \in D_w$. By induction hypothesis $w, g^{x^\tau \mapsto d} \circ \tau \not\Vdash B$. We have $g \circ \tau^{x \mapsto d}(x) = (g^{x^\tau \mapsto d} \circ \tau)(x)$. For any variable $z \neq x$, we also have $g \circ \tau^{x \mapsto d}(z) = (g^{x^\tau \mapsto d} \circ \tau)(z)$ (because $z^\tau \neq x^\tau$, by injectivity of τ). So $g \circ \tau^{x \mapsto d} = g^{x^\tau \mapsto d} \circ \tau$. And so $\mathfrak{M}, w, g \circ \tau^{x \mapsto d} \not\Vdash B$. By definition 2.9, this means that $\mathfrak{M}, w, g \circ \tau \not\Vdash \forall x B$.

6. A is $\Box B$. Assume $\mathfrak{M}, w, g \not\Vdash \Box\tau(B)$. Then $\mathfrak{M}, w', g' \not\Vdash \tau(B)$ for some w', g' with $w, g \rhd w', g'$. This means that there is a counterpart relation $C \in K_{w,w'}$ such that for all variables x, $g'(x)$ is some C-counterpart at w' of $g(x)$ at w (if there is one, else undefined). By induction hypothesis, $w', g' \circ \tau \not\Vdash B$. Since $(g' \circ \tau)(x) = g'(x^\tau)$ for all x, $(g' \circ \tau)(x)$ is a C-counterpart of $(g \circ \tau)(x)$ (if there is one, else undefined), for all x. So $w, (g \circ \tau) \rhd w', (g' \circ \tau)$. Hence $\mathfrak{M}, w', g^* \not\Vdash B$ for some w', g^* with $w, (g \circ \tau) \rhd w', g^*$. (Namely, $g^* = g' \circ \tau$.) So $\mathfrak{M}, w, g \circ \tau \not\Vdash \Box B$.

   In the other direction, assume $\mathfrak{M}, w, g \circ \tau \not\Vdash \Box B$. Then $\mathfrak{M}, w', g^* \not\Vdash B$ for some w', g^* with $w, (g \circ \tau) \rhd w', g^*$. This means that there is a counterpart relation $C \in K_{w,w'}$ such that for all variables x, $g^*(x)$ is some C-counterpart at w' of $g \circ \tau(x)$ at w (if there is one, else undefined). Define g' so that for any variable x, $g'(x^\tau) = g^*(x)$, and for all $x \notin \text{Ran}(\tau)$, $g'(x)$ is an arbitrary C-counterpart of $V_w(x)$, or undefined if there is none. Then $w, g \rhd w', g'$. Moreover, g^* is $g' \circ \tau$. Since $\mathfrak{M}, w', g^* \not\Vdash B$, we have $\mathfrak{M}, w', g' \circ \tau \not\Vdash B$. By induction hypothesis, $\mathfrak{M}, w', g' \not\Vdash \tau(B)$. So $\mathfrak{M}, w', g' \not\Vdash \tau(B)$ for some w', g' with wRw' and $w, g \rhd w', g'$. So $\mathfrak{M}, w, g \not\Vdash \tau(\Box B)$.

There's more trouble ahead. Assume w extends $\{x = y, \Box F x, \Box \neg F y\}$. Let w', τ be such that $\{x^\tau : \Box X \in w\} \subseteq w'$, so that wRw'. Concretely, assume that w' extends $\{Fu, \neg Fv\}$, and $x^\tau = u, y^\tau = v$. So $[x]_w = \{x, y\}$ has two counterparts $\{u\}, \{v\}$ at w', relative to the same counterpart relation C determined by the condition that $[v_i]_w C [u_i]_{w'}$ iff $\tau(v_i) \in [u_i]_{w'}$. For $\Box\phi(x)$ to be true at w, ϕ must be true of *both* counterparts at w'. That is, an assignment g' on $U_{w'}$ that assigns to each variable

a C-counterpart of its g_w-value does not always map v_i to $[\tau(v_i)]_{w'} = g_{w'}(\tau(v_i))$: we can't assume that if $w, g \triangleright w', g'$ then $g' = g_{w'} \circ \tau$.

This not only complicates the proof of the truth lemma, it actually blocks it. To see the problem, assume that the above world w also contains $\Box \Diamond x \neq y$. So

$$w = \{x = y, \Box Fx, \Box \neg Fy, \Box \Diamond x \neq y, \ldots\}.$$

If $w \overset{\tau}{\to} w'$ then w' contains Fx^τ, $\neg Fy^\tau$, and $\Diamond x^\tau \neq y^\tau$. Letting $x^\tau = u$ and $y^\tau = v$, we have

$$w' = \{Fu, \neg Fv, \Diamond u \neq v, \ldots\}.$$

To ensure that $w, g_w \Vdash \Box \Diamond x \neq y$, we must have $w', g' \Vdash \Diamond x \neq y$ for all g' that map individuals in U_w to their C-counterparts in $U_{w'}$, One such g' is $g_{w'} \circ \tau$, but another assigns $[u]_{w'}$ to both x and y. Focus on this one. To ensure that $w', g' \Vdash \Diamond x \neq y$, there must be some w'', g'' such that $w', g' \triangleright w'', g''$ and $w'', g'' \Vdash x \neq y$. So we need a transformation σ, world w'' and assignment g'' such that $\{\sigma(X) : \Box X \in w'\} \subseteq w''$, and for all v_i there is a $v \in g'(v_i)$ such that $\sigma(v) \in g''(v_i)$, and $w'', g'' \Vdash x \neq y$. Since $g'(x) = g'(y) = [u]_{w'}$, this means that $[u]_{w'}$ must have two counterparts at some w'' relative to the same transformation σ. We have no guarantee that this is the case. There has to be a variable z other than u for which w' contains $z = u$ as well as $\Diamond z \neq u$. The latter ensures that $z^\sigma \neq u^\sigma \in w''$ for some w'' with $\{\sigma(X) : \Box X \in w'\} \subseteq w''$.

So we complicate the definition of accessibility in canonical models. We stipulate that if w' does not contain $z = y^\tau$ and $\Diamond z \neq y^\tau$ for some suitable z, then w' is not τ-accessible from w. In general, if w contains $\Box A$ as well as $x = y$, and x is free in A, then for w' to be accessible from w via τ, we require that w' must contain not only A^τ, but also $z = y^\tau$ and $[z/x^\tau]A^\tau$, for some z not free in A^τ.

This requirement is easier to understand if we consider the same situation in a language with substitution. Here $\Box \Diamond x \neq y$ and $x = y$ entail $\Box \langle y : x \rangle \Diamond x \neq y$ (by (LL$_s$) and (S$\Box$)). By the original, simple definition of $w \overset{\tau}{\to} w'$, each world w' accessible from w via τ must contain

$$\langle y^\tau : x^\tau \rangle \Diamond x^\tau \neq y^\tau. \tag{1}$$

This formula expresses that the individual picked out by y^τ has multiple counterparts at some accessible world. Before we worry about images other than g^τ, we ought to make sure that (1) is true at w' under g^τ. This requires that there is a variable z other than y^τ such that w' contains $z = y^\tau$ as well as $\Diamond z \neq y^\tau$. In effect, z is a kind of

witness for the substitution formula (1). Just as an existential formula $\exists x A$ must be witnessed by an instance $[z/x]A$, a substitution formula $\langle y : x \rangle A$ must be witnessed by $[z/x]A$ together with $z = y$. Loosely speaking, $\langle y : x \rangle A(x)$ says that y is identical to some x such that $A(x)$. In a canonical model, we want a concrete witness z so that y is identical to z and $A(z)$. y itself may not serve that purpose, because $\langle y : x \rangle A(x)$ does not guarantee $A(y)$.

The requirement of substitutional witnessing entails that if w contains $\Box A$, then any τ-accessible w' contains not only A^τ, but also $z = y^\tau$ and $[z/x^\tau]A^\tau$ (for some suitable z). So we don't need to complicate the accessibility relation. In our example, since w' contains A^τ whenever w contains $\Box A$, w' contains $\langle y^\tau : x^\tau \rangle \Diamond x^\tau \neq y^\tau$, which settles that $[y^\tau]_{w'}$ has two counterparts at some accessible world. Without substitution, (1) is inexpressible, as per lemma 7.3. So we have to limit the accessible worlds by requiring membership of the relevant witnessing formulas in addition to A^τ.

9.2 Constructing canonical models

From now on, let $\mathfrak{L}$ be some language with or without substitution and let L any positive or negative logic. As in section 4.2, the worlds of the L's canonical model are constructed in a language $\mathfrak{L}^*$ that adds infinitely many new variables Var^+ to those of $\mathfrak{L}$. (Var^* is the set of $\mathfrak{L}^*$-variables.)

We redefine the concept of a Henkin set by adding the condition of substitutional witnessing. The added fourth clause is vacuous if $\mathfrak{L}$ is without substitution.

Definition 9.1 (Henkin set).
A *Henkin set* for L in $\mathfrak{L}^*$ is a set H of $\mathfrak{L}^*$-formulas that is

1. *L-consistent*: there are no $A_1, \dots, A_n \in H$ with $\vdash_L \neg(A_1 \wedge \dots \wedge A_n)$,

2. *maximal*: for every $\mathfrak{L}^*$-formula A, H contains either A or $\neg A$,

3. *witnessed*: whenever H contains an existential formula $\exists x A$, then there is a variable $y \notin \mathrm{FV}(A)$ such that H contains $[y/x]A$ as well as $E!y$, and

4. *substitutionally witnessed*: whenever H contains a substitution formula

---

$\langle y : x \rangle A$ as well as $y = y$, then there is a variable $z \notin \mathrm{Var}(\langle y : x \rangle A)$ such that H contains $y = z$.

As before, $\mathfrak{H}_L$ is the class of Henkin sets for L in $\mathfrak{L}^*$.

Above I said that witnessing a substitution formula $\langle y : x \rangle A$ requires $y = z$ as well as $[z/x]A$. In fact $y = z$ is enough, since $[z/x]A$ follows from $\langle y : x \rangle A$ and $y = z$ by (LV2) (lemma 8.13). I have also added the condition that H contains $y = y$. In negative logics, a Henkin set may contain $y \neq y$ as well as $\langle y : x \rangle A$; adding $y = z$ would render the set inconsistent.

The requirement of substitutional witnessing generalises to substitution sequences:

Lemma 9.2.

If H contains a substitution formula $\langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A$ as well as $y_i = y_i$ for all y_i in $y_1, \ldots, y_n$, then there are (distinct) new variables $z_1, \ldots, z_n$ such that H contains $y_1 = z_1, \ldots, y_n = z_n$ as well as $[z_1, \ldots, z_n / x_1, \ldots, x_n]A$.

Proof. By induction on n. Suppose H contains $\langle y_1, \ldots, y_n : x_1, \ldots, x_n \rangle A$. By definition 8.3, this is $\langle y_n : v \rangle \langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle \langle v : x_n \rangle A$, where v is new. Witnessing requires $y_n = z_n \in H$ and (hence) $[z_n/v]\langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle$ $\langle v : x_n \rangle A = \langle y_1, \ldots, y_{n-1} : x_1, \ldots, x_{n-1} \rangle \langle z_n : x_n \rangle A \in H$ for some new z_n. By induction hypothesis, the latter means that there are (distinct) $z_1, \ldots, z_{n-1} \notin \mathrm{Var}(\langle z_n : x_n \rangle A)$ such that H contains $y_1 = z_1, \ldots, y_{n-1} = z_{n-1}$ as well as $[z_1, \ldots, z_{n-1} / x_1, \ldots, x_{n-1}]\langle z_n : x_n \rangle A$. Since all the x_i and z_i are distinct, $[z_1, \ldots, z_{n-1} / x_1, \ldots, x_{n-1}]$ $\langle z_n : x_n \rangle A$ is $\langle z_n : x_n \rangle [z_1, \ldots, z_{n-1} / x_1, \ldots, x_{n-1}]A$. By (SC1), it follows that $[z_n/x_n][z_1, \ldots, z_{n-1} / x_1, \ldots, x_{n-1}]A = [z_1, \ldots, z_n / x_1, \ldots, x_n]A \in H$.

Next, we define the variable classes $[x]_w$, just as in section 4.4.

Definition 9.2 (Variable classes).

For any Henkin set H, define $\sim_H$ to be the relation on Var^* such that $x \sim_H y$ iff $x = y \in H$. For any variable x, let $[x]_H$ be $\{y : x \sim_H y\}$.

This definition is justified by the following lemma.

Lemma 9.3 (~-Lemma).
$\sim_H$ *is an equivalence relation on the set* $\{x : x = x \in H\}$.

Proof. Immediate from lemmas 7.12 and 8.12.

Now comes the revised definition of $w \xrightarrow{\sigma} w'$.

Definition 9.3 (Accessibility via transformations).
Let w, w' be Henkin sets and τ a transformation.

If $\mathcal{L}$ is with substitution, then w' *is accessible from w via τ* (for short: $w \xrightarrow{\tau} w'$) iff $\{\sigma(X) : \Box X \in w\} \subseteq w'$.

If $\mathcal{L}$ is without substitution, then $w \xrightarrow{\tau} w'$ iff the following holds for every formula A and variables $x_1 \ldots x_n, y_1, \ldots, y_n$ ($n \geq 0$) such that the $x_1 \ldots x_n$ are distinct members of $FV(A)$: if

$$(x_1 = y_1 \wedge \ldots \wedge x_n = y_n \wedge \Box A) \in w$$

and

$$(y_1^\tau = y_1^\tau \wedge \ldots \wedge y_n^\tau = y_n^\tau) \in w',$$

then there are variables $z_1 \ldots z_n \notin \mathrm{Var}(A^\tau)$ such that

$$(z_1 = y_1^\tau \wedge \ldots \wedge z_n = y_n^\tau \wedge [z_1 \ldots z_n / x_1^\tau \ldots x_n^\tau] A^\tau) \in w'.$$

This generalises the witnessing requirements on accessible world as explained above to multiple variables and negative logics. (Here the generalised version for n variable pairs is not entailed by the requirement for a single pair, unlike in the case of substitutional witnessing.) Note that the $x_1, \ldots, x_n$ need not be *all* the free variables in A.

Definition 9.3 is supposed to include the case where $n = 0$. Here, we adopt the convention that a conjunction of zero sentences is the tautology $\top$; the accessibility requirement therefore says that if $\top \wedge \Box A \in w$ and $\top \in w'$, then $\top \wedge A^\tau \in w'$ — equivalently: if $\Box A \in w$, then $A^\tau \in w'$.

Definition 9.4 (Canonical model).

The *canonical model* $\langle W, R, U, D, K, I \rangle$ for a logic L is defined as follows.

1. The *worlds* W are the Henkin sets $\mathfrak{H}_L$.

2. For each $w \in W$, the *outer domain* U_w comprises all non-empty sets $[x]_w$ for x in Var^*.

3. For each $w \in W$, the *inner domain* D_w comprises all sets $[x]_w$ for which $E!x \in w$. That is, $D_w = \{[x]_w : E!x \in w\}$.

4. The *accessibility relation* R holds between world w and world w' iff there is a transformation τ such that $w \xrightarrow{\tau} w'$.

5. C is a *counterpart relation* in $K_{w,w'}$ iff there is a transformation τ such that (i) $w \xrightarrow{\tau} w'$ and (ii) for all $d \in U_w, d' \in U_{w'}$, dCd' iff there is an $x \in d$ with $\tau(x) \in d'$.

6. The *interpretation* I assigns to every non-logical predicate P and world w the set $I_w(P) = \{\langle [x_1]_w, \dots, [x_n]_w \rangle : Px_1 \dots x_n \in w\}$.

Definition 9.5 (Canonical Assignment).

If w is a world in a canonical model $\mathfrak{M}$ then the *canonical assignment for w* is the function g_w that maps every $x \in \mathrm{Var}^*$ for which $[x]_w$ is non-empty to $[x]_w$.

Lemma 9.4 (Charge of canonical models).

If L is positive then the canonical model for L is total. If L is negative then the canonical model for L is single-domain.

Proof. The proof is exactly as in lemma 4.3.

Lemma 9.5 (Extensibility Lemma).

If Γ is an L-consistent set of $\mathfrak{L}^*$-sentences in which infinitely many $\mathfrak{L}^*$-

variables do not occur, then there is a Henkin set $H \in \mathfrak{H}_L$ such that $\Gamma \subseteq H$.

Proof. Let $S_1, S_2, \ldots$ be an enumeration of all $\mathfrak{L}^*$-sentences, and $z_1, z_2, \ldots$ an enumeration of the unused $\mathfrak{L}^*$-variables in such a way that $z_i \notin \mathrm{Var}(S_1 \wedge \ldots \wedge S_i)$. Let $\Gamma_0 = \Gamma$, and define Γ_n for $n \geq 1$ as follows.

(i) If $\Gamma_{n-1} \cup \{S_n\}$ is not L-consistent, then $\Gamma_n = \Gamma_{n-1}$;

(ii) else if S_n is an existential formula $\exists x A$, then $\Gamma_n = \Gamma_{n-1} \cup \{\exists x A, [z_n/x]A, E!z_n\}$;

(iii) else if S_n is a substitution formula $\langle y : x \rangle A$, then $\Gamma_n = \Gamma_{n-1} \cup \{\langle y : x \rangle A, y = y \supset y = z_n\}$;

(iv) else $\Gamma_n = \Gamma_{n-1} \cup \{S_n\}$.

Define w as the union of all Γ_n. We show that w is a Henkin set for L.

1. w is L-consistent. This is shown by proving that Γ_0 is L-consistent and that whenever Γ_{n-1} is L-consistent, then so is Γ_n. It follows that no finite subset of w is L-inconsistent, and hence that w itself is L-consistent. The base step, that Γ_0 is L-consistent is given by assumption. Now assume (for $n > 0$) that Γ_{n-1} is L-consistent. Then Γ_n is constructed by applying one of (i)–(iv).

   a) If case (i) in the construction applies, then $\Gamma_n = \Gamma_{n-1}$, and so Γ_n is also L-consistent.

   b) Assume case (ii) in the construction applies, and suppose that $\Gamma_n = \Gamma_{n-1} \cup \{\exists x A, [z_n/x]A, E!z\}$ is L-inconsistent. Then there is a finite subset $\{C_1, \ldots, C_m\} \subseteq \Gamma_{n-1}$ such that

   $$1. \quad \vdash_L \neg(C_1 \wedge \ldots \wedge C_m \wedge \exists x A \wedge [z_n/x]A \wedge E!z_n).$$

Let $\underline{C}$ abbreviate $C_1 \wedge \ldots \wedge C_m$. Then

2. $\vdash_L \underline{C} \wedge \exists x A \supset (E!z_n \supset \neg[z_n/x]A)$ (1)

3. $\vdash_L \forall z_n(\underline{C} \wedge \exists x A) \supset \forall z_n E!z_n \supset \forall z_n \neg[z_n/x]A$ (2, (UG), (UD))

4. $\vdash_L \underline{C} \wedge \exists x A \supset \forall z_n(\underline{C} \wedge \exists x A)$ ((VQ), z_n not in Γ_{n-1})

5. $\vdash_L \underline{C} \wedge \exists x A \supset \forall z_n E!z_n \supset \forall z_n \neg[z_n/x]A.$ (3, 4)

6. $\vdash_L \underline{C} \wedge \exists x A \supset \forall z_n \neg[z_n/x]A.$ (5, ($\forall$E!x))

7. $\vdash_L \forall z_n \neg[z_n/x]A \leftrightarrow \forall x \neg A$ ((AC), $z_n \notin \mathrm{Var}(A)$)

8. $\vdash_L \underline{C} \wedge \exists x A \supset \neg \exists x A.$ (6, 7)

So $\{C_1, \ldots C_m, \exists x A\}$ is not L-consistent, contradicting the assumption that clause (ii) applies.

c) Assume case (iii) in the construction applies (hence L is with substitution), and suppose that $\Gamma_n = \Gamma_{n-1} \cup \{\langle y : x \rangle A, y = y \supset y = z_n\}$ is L-inconsistent. Then there is a finite subset $\{C_1, \ldots, C_m\} \subseteq \Gamma_{n-1}$ such that

1. $\vdash_L \neg(\underline{C} \wedge \langle y : x \rangle A \wedge (y = y \supset y \neq z)).$

(As before, $\underline{C}$ is $C_1 \wedge \ldots \wedge C_m$.) But then

2. $\vdash_L \underline{C} \wedge \langle y : x \rangle A \supset y = y \wedge y \neq z_n$ (1)

3. $\vdash_L \langle y : z_n \rangle(\underline{C} \wedge \langle y : x \rangle A \supset y = y \wedge y \neq z_n)$ (2, (Sub$_s$))

4. $\vdash_L \langle y : z_n \rangle(\underline{C} \wedge \langle y : x \rangle A) \supset \langle y : z_n \rangle y = y \wedge \langle y : z_n \rangle y \neq z_n$ (3, (S$\supset$), (S$\neg$))

5. $\vdash_L \underline{C} \wedge \langle y : x \rangle A \supset \langle y : z_n \rangle(\underline{C} \wedge \langle y : x \rangle A)$ ((VS), z_n not in Γ_{n-1}, S_n)

6. $\vdash_L \underline{C} \wedge \langle y : x \rangle A \supset \langle y : z_n \rangle y = y \wedge \langle y : z_n \rangle y \neq z_n$ (4, 5)

7. $\vdash_L \langle y : z_n \rangle y \neq z_n \leftrightarrow y \neq y$ (SAt)

8. $\vdash_L \langle y : z_n \rangle y = y \leftrightarrow y = y$ (SAt)

9. $\vdash_L \underline{C} \wedge \langle y : x \rangle A \supset (y = y \wedge y \neq y).$ (6, 7, 8)

So $\{C_1, \ldots, C_m, \langle y : x \rangle A\}$ is L-inconsistent, contradicting the assumption

that clause (iii) applies.

d) Assume case (iv) in the construction applies. Then $\Gamma_n = \Gamma_{n-1} \cup \{S_n\}$ is L-consistent, since otherwise (i) would have applied.

2. w is maximal. Assume some formula S_n is not in w. Then case (i) applied to S_n, so $\Gamma_{n-1} \cup \{S_n\}$ is not L-consistent. So there are $C_1, \dots, C_m \in \Gamma_{n-1}$ such that $\vdash_L C_1 \wedge \dots C_m \supset \neg S_n$. Similarly, if $S_k = \neg S_n$ is not in w, then there are $D_1, \dots, D_l \in \Gamma_{k-1}$ such that $\vdash_L D_1 \wedge \dots D_l \supset \neg S_k$. By (PC), it follows that there are $C_1, \dots, C_m, D_1, \dots D_l \in w$ such that

$$\vdash_L C_1 \wedge \dots \wedge C_m \wedge D_1 \wedge \dots \wedge D_l \supset (\neg S_n \wedge \neg\neg S_n).$$

But then w is inconsistent, contradicting what was just shown under 1.

3. w is witnessed. This is guaranteed by clause (ii) of the construction and the fact that the $z_n \notin \mathrm{Var}(S_n)$.

4. w is substitutionally witnessed. This is guaranteed by clause (iii) and the fact that the $z_n \notin \mathrm{Var}(S_n)$.

Lemma 9.6 (Existence Lemma).
If w is a world in the canonical model for L, A a formula with $\Diamond A \in w$, and τ any transformation whose range excludes infinitely many variables of $\mathfrak{L}$, then there is a world w' in the model such that $w \xrightarrow{\tau} w'$ and $A^\tau \in w'$.

Proof. I first prove the lemma for logics L with substitution. Let $\Gamma = \{A^\tau\} \cup \{\tau(B) : \Box B \in w\}$. Suppose Γ is not L-consistent. Then there are $B_1^\tau, \dots, B_n^\tau$ with $\Box B_i \in w$ such that $\vdash_L B_1^\tau \wedge \dots \wedge B_n^\tau \supset \neg A^\tau$. By definition 3.1, this means that $\vdash_L (B_1 \wedge \dots B_n \supset \neg A)^\tau$, and so $\vdash_L B_1 \wedge \dots \wedge B_n \supset \neg A$ by (Sub^t). By (Nec) and (K), $\vdash_L \Box B_1 \wedge \dots \wedge \Box B_n \supset \Box \neg A$. But then w contains both $\Diamond A$ and $\neg \Diamond A$, which is impossible because w is L-consistent. So Γ is L-consistent. Since the range of τ excludes infinitely many variables, it follows by the extensibility

lemma that $\Gamma \subseteq H$ for some Henkin set H. Moreover, $w \xrightarrow{\tau} w'$ because $\tau(B) \in H$ whenever for $\Box B \in w$.

Now for logics without substitution.

Let $S_1, S_2 \ldots$ enumerate all sentences in w of the form

$$x_1 = y_1 \land \ldots \land x_n = y_n \land \Box B,$$

where $x_1, \ldots, x_n$ are zero or more distinct variables free in B. Let U be the "unused" $\mathfrak{L}$-variables that are not in the range of τ. Let Z be an infinite subset of U such that $Z \backslash U$ is also infinite. For each $S_i = (x_1 = y_1 \land \ldots \land x_n = y_n \land \Box B)$, let Z_{S_i} be a set of distinct variables $z_1, \ldots, z_n \in Z$ such that $Z_{S_i} \cap \bigcup_{j<i} Z_{S_j} = \emptyset$ (i.e. none of the z_i has been used for any earlier S_j). Abbreviate

$$B_i = [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B);$$
$$X_i = x_1 = y_1 \land \ldots \land x_n = y_n;$$
$$Y_i = y_1^\tau = y_1^\tau \land \ldots \land y_n^\tau = y_n^\tau;$$
$$Z_i = y_1^\tau = z_1 \land \ldots \land y_n^\tau = z_n.$$

(For $n = 0$, X_i, Y_i and Z_i are the tautology $\top$, and B_i is $\tau(B)$.)

Let $\Gamma^- = \{(Y_i \supset Z_i \land B_i) : S_i \in S_1, S_2, \ldots\}$, and let $\Gamma = \Gamma^- \cup \{A^\tau\}$.

Suppose for reductio that Γ is inconsistent. Then there are sentences $(Y_1 \supset Z_1 \land B_1), \ldots, (Y_m \supset Z_m \land B_m) \in \Gamma^-$ such that

$$\vdash_L \neg(A^\tau \land (Y_1 \supset Z_1 \land B_1) \land \ldots \land (Y_m \supset Z_m \land B_m)). \tag{1}$$

By (Nec),

$$\vdash_L \Box\neg(A^\tau \land (Y_1 \supset Z_1 \land B_1) \land \ldots \land (Y_m \supset Z_m \land B_m)). \tag{2}$$

Any member $(Y_i \supset Z_i \land B_i)$ of Γ^- has the form

$$y_1^\tau = y_1^\tau \land \ldots \land y_n^\tau = y_n^\tau \supset y_1^\tau = z_1 \land \ldots \land y_n^\tau = z_n \land [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B).$$

By (CS_n),

$$\vdash_L x_1^\tau = y_1^\tau \wedge ... \wedge x_n^\tau = y_n^\tau \wedge \Box \tau(B) \supset$$
$$\Box(y_1^\tau = z_1 \wedge ... \wedge y_n^\tau = z_n \supset [z_1, ..., z_n / x_1^\tau, ..., x_n^\tau] \tau(B)). \quad (3)$$

Now w contains $x_1 = y_1 \wedge ... \wedge x_n = y_n \wedge \Box B$. So w^τ contains $x_1^\tau = y_1^\tau \wedge ... \wedge x_n^\tau = y_n^\tau \wedge \Box \tau(B)$, which is the antecedent of (3). The consequent of (3) is $\Box(Z_i \supset B_i)$. Thus

$$w^\tau \vdash_L \Box(Z_1 \supset B_1) \wedge ... \wedge \Box(Z_m \supset B_m). \quad (4)$$

(By $\Gamma \vdash_L A$, I mean that there are sentences $A_1, ..., A_n \in \Gamma$ such that $\vdash_L A_1 \wedge ... \wedge A_n \supset A$.) Let $\Delta = w^\tau \cup \{\Diamond(A^\tau \wedge (Y_1 \supset Z_1) \wedge ... \wedge (Y_m \supset Z_m))\}$. So

$$\Delta \vdash_L \Box(Z_1 \supset B_1) \wedge ... \wedge \Box(Z_m \supset B_m); \quad (5)$$
$$\Delta \vdash_L \Diamond(A^\tau \wedge (Y_1 \supset Z_1) \wedge ... \wedge (Y_m \supset Z_m)). \quad (6)$$

By (K) and (Nec), (5) and (6) yield

$$\Delta \vdash_L \Diamond(A^\tau \wedge (Y_1 \supset Z_1 \wedge B_1) \wedge ... \wedge (Y_m \supset Z_m \wedge B_m)). \quad (7)$$

By (2), it follows that Δ is inconsistent. This means that

$$w^\tau \vdash_L \neg \Diamond(A^\tau \wedge (Y_1 \supset Z_1) \wedge ... \wedge (Y_m \supset Z_m)). \quad (8)$$

Now consider $Z_1 = (y_1^\tau = z_1 \wedge ... \wedge y_n^\tau = z_n)$. By (LL_n^*) (or repeated application of (LL^*)),

$$\vdash_L y_1^\tau = z_1 \wedge ... \wedge y_n^\tau = z_n \supset \Box \neg(A^\tau \wedge (y_1^\tau = y_1^\tau \wedge ... \wedge y_n^\tau = y_n^\tau \supset y_1^\tau = z_1 \wedge ... \wedge y_n^\tau = z_n))$$
$$\supset \Box \neg(A^\tau \wedge (y_1^\tau = y_1^\tau \wedge ... \wedge y_n^\tau = y_n^\tau \supset y_1^\tau = y_1^\tau \wedge ... \wedge y_n^\tau = y_n^\tau)), \quad (9)$$

because the z_i are not free in A^τ. In other words (and dropping the tautologous conjunct at the end),

$$\vdash_L Z_1 \supset \Box \neg(A^\tau \wedge (Y_1 \supset Z_1)) \supset \Box \neg A^\tau. \quad (10)$$

By the same reasoning,

$$\vdash_L Z_1 \wedge ... \wedge Z_m \supset \Box\neg(A^\tau \wedge (Y_1 \supset Z_1) \wedge ... \wedge (Y_m \supset Z_m)) \supset \Box\neg A^\tau. \quad (11)$$

By (PC), (Nec) and (K), this means

$$\vdash_L Z_1 \wedge ... \wedge Z_m \supset \Diamond A^\tau \supset \Diamond(A^\tau \wedge (Y_1 \supset Z_1) \wedge ... \wedge (Y_m \supset Z_m)). \quad (12)$$

Since $w^\tau \vdash_L \Diamond A^\tau$, (8) and (12) together entail

$$w^\tau \vdash_L \neg(Z_1 \wedge ... \wedge Z_m). \quad (13)$$

So there are $C_1, ..., C_k \in w$ such that

$$\vdash_L C_1^\tau \wedge ... \wedge C_k^\tau \supset \neg(Z_1 \wedge ... \wedge Z_m). \quad (14)$$

Each Z_i has the form $y_1^\tau = z_1 \wedge ... \wedge y_n^\tau = z_n$. All the z_i are distinct, and none of them occur in $C_1^\tau \wedge ... \wedge C_k^\tau$ (because the z_i are not in the range of τ) nor in any other Z_i. By (Sub*), we can therefore replace each z_i in (14) by the corresponding y_i^τ, turning Z_i into Y_i:

$$\vdash_L C_1^\tau \wedge ... \wedge C_k^\tau \supset \neg(Y_1 \wedge ... \wedge Y_m). \quad (15)$$

For any $Y_i = (y_1^\tau = y_1^\tau \wedge ... \wedge y_n^\tau = y_n^\tau)$, X_i is a sentence of the form $x_1 = y_1 \wedge ... \wedge x_n = y_n$. So X_i^τ is $x_1^\tau = y_1^\tau \wedge ... \wedge x_n^\tau = y_n^\tau$, and $\vdash_L X_i^\tau \supset Y_i$ by either (= R) or (Neg) and ($\forall$ = R). So (15) entails

$$\vdash_L C_1^\tau \wedge ... \wedge C_k^\tau \supset \neg(X_1^\tau \wedge ... \wedge X_m^\tau). \quad (16)$$

Thus by (Sub^t),

$$\vdash_L C_1 \wedge ... \wedge C_k \supset \neg(X_1 \wedge ... \wedge X_m). \quad (17)$$

Since $\{C_1, ..., C_k, X_1, ..., X_m\} \subseteq w$, it follows that w is inconsistent. Which it isn't. This completes the reductio.

So Γ is consistent. Since the infinitely many variables in $U \backslash Z$ do not occur

in Γ, lemma 9.5 guarantees that $\Gamma \subseteq w'$ for some world w' in the canonical model for L. And of course, Γ was constructed so that w' satisfies the condition in definition 9.3 for $w \xrightarrow{\tau} w'$. This requires that for every formula B and variables $x_1 \ldots x_n, y_1, \ldots, y_n$ such that the $x_1 \ldots x_n$ are zero or more pairwise distinct members of $FV(B)$, if $x_1 = y_1 \wedge \ldots \wedge x_n = y_n \wedge \Box B \in w$ and $y_1^\tau = y_1^\tau \wedge \ldots \wedge y_n^\tau = y_n^\tau \in w'$, then there are variables $z_1 \ldots z_n \notin \mathrm{Var}(\tau(B))$ such that $z_1 = y_1^\tau \wedge \ldots \wedge z_n = y_n^\tau \wedge [z_1 \ldots z_n / x_1^\tau \ldots x_n^\tau] \tau(B) \in w'$. By construction of Γ, whenever $x_1 = y_1 \wedge \ldots \wedge x_n = y_n \wedge \Box B \in w$, then there are suitable $z_1, \ldots, z_n$ such that $y_1^\tau = y_1^\tau \wedge \ldots \wedge y_n^\tau = y_n^\tau \supset y_1^\tau = z_1 \wedge \ldots \wedge y_n^\tau = z_n \wedge [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B) \in w'$. So if $y_1^\tau = y_1^\tau \wedge \ldots \wedge y_n^\tau = y_n^\tau \in w'$, then $y_1^\tau = z_1 \wedge \ldots \wedge y_n^\tau = z_n \wedge [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B) \in w'$.

Lemma 9.7 (Truth Lemma).
If $\mathfrak{M} = \langle W, R, U, D, K, I \rangle$ is the canonical model for a positive or negative logic L, $w \in W$, and g_w is the canonical assignment for w, then for any $\mathfrak{L}$-sentence A,

$$\mathfrak{M}, w, g_w \Vdash A \text{ iff } A \in w.$$

Proof. By induction on A.

1. A is $Px_1 \ldots x_n$.

   By definition 2.9, $\mathfrak{M}, w, g_w \Vdash Px_1 \ldots x_n$ iff $\langle g_w(x_1), \ldots, g_w(x_n) \rangle \in I_w(P)$. By definition 9.5, $g_w(x_i)$ is $[x_i]_w$ or undefined if $[x_i]_w = \emptyset$. Moreover, $I_w(P) = \{\langle [z_1]_w, \ldots, [z_n]_w \rangle : Pz_1 \ldots z_n \in w\}$ by definition 9.4 and, for the identity predicate, by the fact $I_w(=)$ is $\{\langle d, d \rangle : d \in U_w\} = \{\langle [z]_w, [z]_w \rangle : z = z \in w\} = \{\langle [z_1]_w, [z_2]_w \rangle : z_1 = z_2 \in w\}$.

   Now if $\langle g_w(x_1), \ldots, g_w(x_n) \rangle \in I_w(P)$, then $\langle [x_1]_w, \ldots, [x_n]_w \rangle \in \{\langle [z_1]_w, \ldots, [z_n]_w \rangle : Pz_1 \ldots z_n \in w\}$, where all the $[x_i]_w$ are non-empty (for $g_w(x_i)$ is defined). This means that there are variables $z_1, \ldots, z_n$ such that $\{x_1 = z_1, \ldots, x_n = z_n, Pz_1 \ldots z_n\} \subseteq w$. Then $Px_1 \ldots x_n \in w$ by (LL*).

   In the other direction, if $Px_1 \ldots x_n \in w$ then $x_i = x_i \in w$ for all x_i in

$x_1 \ldots x_n$. Hence $\langle [x_1]_w, \ldots, [x_n]_w \rangle \in \{\langle [z_1]_w, \ldots, [z_n]_w \rangle : Pz_1 \ldots z_n \in w\}$, and $\langle g_w(x_1), \ldots, g_w(x_n) \rangle \in I_w(P)$.

2. A is $\neg B$.

   $\mathfrak{M}, w, g_w \Vdash \neg B$ iff $\mathfrak{M}, w, g_w \nVdash B$ by definition 2.9, iff $B \notin w$ by induction hypothesis, iff $\neg B \in w$ by definition 9.1.

3. A is $B \supset C$.

   $\mathfrak{M}, w, g_w \Vdash B \supset C$ iff $\mathfrak{M}, w, g_w \nVdash B$ or $\mathfrak{M}, w, g_w \Vdash C$ by definition 2.9, iff $B \notin w$ or $C \in w$ by induction hypothesis, iff $B \supset C \in w$ by definition 9.1 and the fact that $B \supset C$ is L-entailed by $\neg B$ and by C.

4. A is $\langle y : x \rangle B$.

   Assume first that $\mathfrak{M}, w, g_w \Vdash y \neq y$. So $g_w(y)$ is undefined, and it is not the case that $g_w(y)$ has multiple counterparts at any world. Then $\mathfrak{M}, w, g_w \Vdash \langle y : x \rangle B$ iff $\mathfrak{M}, w, g_w^{[y/x]} \Vdash B$ by definition 8.2, iff $\mathfrak{M}, w, g_w \Vdash [y/x]B$ by lemma 7.2, iff $[y/x]B \in w$ by induction hypothesis. Also by induction hypothesis, $y \neq y \in w$. By (SCN), $\vdash_L y \neq y \supset ([y/x]B \leftrightarrow \langle y : x \rangle B)$. So $[y/x]B \in w$ iff $\langle y : x \rangle B \in w$.

   Next, assume that $\mathfrak{M}, w, g_w \Vdash y = y$; so by induction hypothesis $y = y \in w$. Assume further that $\langle y : x \rangle B \notin w$. Then $\neg \langle y : x \rangle B \in w$ by maximality of w, and $\langle y : x \rangle \neg B \in w$ by (S$\neg$). Since w is substitutionally witnessed and $y = y \in w$, there is a variable $z \notin \mathrm{Var}(\langle y : x \rangle \neg B)$ such that $y = z \in w$ and $[z/x]\neg B \in w$. By induction hypothesis, $\mathfrak{M}, w, g_w \Vdash y = z$. Moreover, by definition 3.1, $\neg [z/x]B \in w$, and so $[z/x]B \notin w$ by consistency of w. By induction hypothesis, $\mathfrak{M}, w, g_w \nVdash [z/x]B$. By definition 2.9, then $\mathfrak{M}, w, g_w \Vdash \neg [z/x]B$, i.e. $\mathfrak{M}, w, g_w \Vdash [z/x]\neg B$. Since z and x are modally separated in B, then $\mathfrak{M}, w, g_w^{[z/x]} \Vdash \neg B$ by lemma 7.2. But $g_w^{[z/x]} = g_w^{[y/x]}$ because $\mathfrak{M}, w, g_w \Vdash y = z$. So $\mathfrak{M}, w, g_w^{[y/x]} \Vdash \neg B$. And so $\mathfrak{M}, w, g_w^{[y/x]} \nVdash B$ by definition 2.9, and $\mathfrak{M}, w, g_w \nVdash \langle y : x \rangle B$ by definition 8.2.

   In the other direction, assume $\langle y : x \rangle B \in w$. Since w is substitutionally witnessed and $y = y \in w$, there is a new variable z such that $y = z \in w$ and $[z/x]B \in w$. By induction hypothesis, $\mathfrak{M}, w, g_w \Vdash y = z$ and $\mathfrak{M}, w, g_w \Vdash$

$[z/x]B$. Since z and x are modally separated in B, $\mathfrak{M}, w, g_w^{[z/x]} \Vdash B$ by lemma 7.2. As before, $g_w^{[z/x]} = g_w^{[y/x]}$ because $\mathfrak{M}, w, g_w \Vdash y = z$; so $\mathfrak{M}, w, g_w^{[y/x]} \Vdash B$, and so $\mathfrak{M}, w, g_w \Vdash \langle y : x \rangle B$ by definition 8.2.

5. A is $\forall x B$.

   We first show that for any variable x, $\mathfrak{M}, w, g_w \Vdash E!x$ iff $E!x \in w$: $\mathfrak{M}, w, g_w \Vdash E!x$ iff $g_w(x) \in D_w$ by definition 3.1, iff $[x]_w \in D_w$ by definition 9.4, iff $E!x \in w$ by definition 9.4.

   Now assume $\forall x B \in w$, and let y be any variable such that $E!y \in w$. As just shown, $\mathfrak{M}, w, g_w \Vdash E!y$. By (FUI**), $\exists x(x = y \wedge B) \in w$. By witnessing, there is a $z \notin \mathrm{Var}(B)$ such that $z = y \wedge [z/x]B \in w$, and thus $z = y \in w$ and $[z/x]B \in w$. By induction hypothesis, $\mathfrak{M}, w, g_w \Vdash z = y$ and $\mathfrak{M}, w, g_w \Vdash [z/x]B$. By lemma 7.2, then $\mathfrak{M}, w, g_w^{[z/x]} \Vdash B$. And since $g_w(z) = g_w(y)$, we have $\mathfrak{M}, w, g_w^{[y/x]} \Vdash B$. So if $\forall x B \in w$, then $\mathfrak{M}, w, g_w^{[y/x]} \Vdash B$ for all variables y with $E!y \in w$, i.e. with $g_w(y) \in D_w$. Since every member $[y]_w$ of D_w is denoted by some variable y under g_w, this means that $\mathfrak{M}, w, g_w^{x \mapsto d} \Vdash B$ for all $d \in D_w$. So $\mathfrak{M}, w, g_w \Vdash \forall x B$.

   Conversely, assume $\forall x B \notin w$. Then $\exists x \neg B \in w$; so by witnessing, $[y/x]\neg B \in w$ for some $y \notin \mathrm{Var}(B)$ with $E!y \in w$. Then $\neg[y/x]B \in w$ and so $[y/x]B \notin w$. As shown above, $\mathfrak{M}, w, g_w \Vdash E!y$. Moreover, by induction hypothesis, $\mathfrak{M}, w, g_w \nVdash [y/x]B$. By lemma 7.2, then $\mathfrak{M}, w, g_w^{[y/x]} \nVdash B$. And so $\mathfrak{M}, w, g_w^{x \mapsto g_w^{[y/x]}(x)} \nVdash B$. So $\mathfrak{M}, w, g_w \nVdash \forall x B$.

6. A is $\Box B$.

   Assume $\mathfrak{M}, w, g_w \Vdash \Box B$. Then $\mathfrak{M}, w', g_w' \Vdash B$ for all w', g_w' with $w, g_w \rhd w', g_w'$. We first show that if $w \xrightarrow{\tau} w'$ then $w, g_w \rhd w', g_{w'} \circ \tau$. By definitions 2.8 and 9.4, $w, g_w \rhd w', g_{w'} \circ \tau$ means that there is a transformation σ such that $w \xrightarrow{\sigma} w'$ and for every variable y, if there is a $z \in g_w(y)$ such that $[z^\sigma]_{w'} \in U_{w'}$ (i.e., if $g_w(y)$ has any σ-counterpart at w'), then there is a $z \in g_w(y)$ with $z^\sigma \in (g_{w'} \circ \tau)(y)$ (i.e., then $(g_{w'} \circ \tau)(y)$ is such a counterpart), otherwise $(g_{w'} \circ \tau)(y)$ is undefined. The relevant transformation σ will be τ. So what we'll show is this: for every variable y, if there is a $z \in g_w(y)$ such that

145

$[z^\tau]_{w'} \in U_{w'}$, then there is a $z \in g_w(y)$ with $z^\tau \in (g_{w'} \circ \tau)(y)$, otherwise $(g_{w'} \circ \tau)(y)$ is undefined.

Let y be any variable. Assume first that there is a $z \in g_w(y)$ such that $[z^\tau]_{w'} \in U_{w'}$. Then $z = y \in w$ and $z^\tau = z^\tau \in w'$. By either (Neg) and (EI) or (=R), $\vdash_L z = y \supset y = y$; so $y = y \in w$. Moreover, by either (TE), (EI), (Nec) and (K) or (=R) and (Nec), $\vdash_L z = y \supset \Box(z = z \supset y = y)$; so $\Box(z = z \supset y = y) \in w$. By definition of $w \xrightarrow{\tau} w'$, then $z^\tau = z^\tau \supset y^\tau = y^\tau \in w'$. So $y^\tau = y^\tau \in w'$. Hence $y \in g_w(y)$ and $y^\tau \in [y^\tau]_{w'} = g_{w'}(y^\tau) = (g_{w'} \circ \tau)(y)$. Alternatively, assume there is no $z \in g_w(y)$ with $z^\tau = z^\tau \in w'$. Then either $g_w(y) = \emptyset$, in which case $y \neq y \in w$, and so $\Box(y \neq y) \in w$ by (NA), (EI), (Nec) and (K), and $y^\tau \neq y^\tau \in w'$ by definition of $w \xrightarrow{\tau} w'$, or else $g_w(y) \neq \emptyset$, but $z^\tau \neq z^\tau \in w'$ for all $z \in g_w(y)$, in which case, too, $y^\tau \neq y^\tau \in w'$ since $y \in g_w(y)$. Either way, $g_{w'}(y^\tau) = (g_{w'} \circ \tau)(y)$ is undefined.

We've shown that if $\mathfrak{M}, w, g_w \Vdash \Box B$, then for every w' and τ with $w \xrightarrow{\tau} w'$, $\mathfrak{M}, w', g^\tau_{w'} \Vdash B$. By the transformation lemma, then $\mathfrak{M}, w', g_{w'} \Vdash \tau(B)$. By induction hypothesis, $B^\tau \in w'$. Now suppose $\Box B \notin w$. Then $\Diamond \neg B \in w$ by maximality of w. By the existence lemma, there is then a world w' and transformation τ with $w \xrightarrow{\tau} w'$ and $\neg B^\tau \in w'$. (Any transformation whose range excludes infinitely many variables will do.) But we've just seen that if $w \xrightarrow{\tau} w'$, then $\tau(B) \in w'$. So if $\mathfrak{M}, w, g_w \Vdash \Box B$, then $\Box B \in w$.

For the other direction, assume $\mathfrak{M}, w, g_w \nVdash \Box B$. So $\mathfrak{M}, w', g' \nVdash B$ for some w', g' with $w, g_w \rhd w', g'$. As before, this means that there is a transformation τ with $w \xrightarrow{\tau} w'$ such that for every variable x, either there is a $y \in g_w(x)$ with $y^\tau \in g'(x)$, or there is no $y \in g_w(x)$ with $y^\tau = y^\tau \in w'$, in which case $g'(x)$ is undefined. Let τ be any transformation with $w \xrightarrow{\tau} w'$, and let $*$ be a substitution that maps each variable x in B to some member y of $g_w(x)$ with $y^\tau \in g'(x)$, or to itself if there is no such y. Thus if $x \in \mathrm{Var}(B)$ and $g'(x)$ is defined, then $(*x)^\tau \in g'(x)$, and so $g'(x) = [(*x)^\tau]_{w'} = (g_{w'} \circ \tau \circ *)(x)$. Alternatively, if $g'(x)$ is undefined (so $*x = x$), then $(g_{w'} \circ \tau \circ *)(x) = g_{w'}(\tau(x))$ is also undefined. That's because otherwise $g_{w'}(\tau(x)) = [x^\tau]_{w'} \neq \emptyset$ and $x^\tau = x^\tau \in w'$; since $w \xrightarrow{\tau} w'$, then $\Box x \neq x \notin w$ and hence $x = x \in w$, as $\vdash_L x \neq x \supset \Box x \neq x$; so there is a $y \in g_w(x)$, namely

x, such that $y^\tau = y^\tau \in w'$, in which case $g'(x)$ cannot be undefined (by definition 2.8). So $g' = g_{w'} \circ \tau \circ *$, and $\mathfrak{M}, w', g_{w'} \circ \tau \circ * \not\Vdash B$.

Now suppose for reductio that $\Box B \in w$. Let $x_1, \ldots, x_n$ be the variables x in $\mathrm{Var}(B)$ with $(*x)^\tau \in g'(x)$ (thus excluding empty variables as well as variables denoting individuals without τ-counterparts at w'). For each such x_i, $*x_i \in g_w(x_i)$, and so $x_i = *x_i \in w$.

If L is with substitution, then $\langle *x_1, \ldots, *x_n : x_1, \ldots, x_n \rangle \Box B \in w$ by (LL$_n$), and so $\Box \langle *x_1, \ldots, *x_n : x_1, \ldots, x_n \rangle B \in w$ by (S$\Box$). Since $w \xrightarrow{\tau} w'$, then $\langle (*x_1)^\tau, \ldots, (*x_n)^\tau : x_1^\tau, \ldots, x_n^\tau \rangle \tau(B) \in w'$. By substitutional witnessing, it follows that there are new variables $z_1, \ldots, z_n$ such that $z_i = (*x_i)^\tau \in w'$ and (hence) $[z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B) \in w'$.

If L is without substitution, this fact – that there are new variables $z_1, \ldots, z_n$ such that $z_i = (*x_i)^\tau \in w'$ and $[z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B) \in w'$ – is guaranteed directly by definition of $w \xrightarrow{\tau} w'$ and the fact that $\Box B \in w$.

By induction hypothesis, $\mathfrak{M}, w', g_{w'} \Vdash z_i = (*x_i)^\tau$ and $\mathfrak{M}, w', g_{w'} \Vdash [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \tau(B)$. Since the z_i are new, $\mathfrak{M}, w', g_{w'} \circ [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \Vdash \tau(B)$ by lemma 7.2. By the transformation lemma, it follows that $\mathfrak{M}, w', g_{w'} \circ [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \circ \tau \Vdash B$. However, for each x_i, $(g_{w'} \circ [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \circ \tau)(x_i) = (g_{w'} \circ [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau])(x_i^\tau) = g_{w'}(z_i) = g_{w'}((*x_i)^\tau)$ [because $\mathfrak{M}, w', g_{w'} \Vdash (*x_i)^\tau = z_i$] $= g_{w'}^\tau(*x_i) = (g_{w'} \circ \tau \circ [*x_1, \ldots, *x_n / x_1, \ldots, x_n])(x_i) = (g_{w'} \circ \tau \circ *)(x_i)$. Similarly, if $x \in \mathrm{Var}(B)$ is none of the $x_1, \ldots, x_n$, so $(*x)^\tau \notin g'(x)$, then $*x$ is x by definition of $*$, and so $(g_{w'} \circ [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \circ \tau)(x) = g_{w'}^\tau(x) = g_{w'}^\tau(*x) = (g_{w'} \circ \tau \circ *)(x)$. So $g_{w'} \circ [z_1, \ldots, z_n / x_1^\tau, \ldots, x_n^\tau] \circ \tau$ coincides with $g_{w'} \circ \tau \circ *$ for all variables in B. By the locality lemma, it follows that $\mathfrak{M}, w', g_{w'} \circ \tau \circ * \Vdash B$ – contradiction.

9.3 Completeness

As in section 4.3, we can use the canonical model technique to prove completeness of our base logics.

Lemma 9.8 (Completeness lemma).
Every positive or negative modal predicate logic is strongly complete with respect to any class of counterpart structures that contains the structure of its canonical model.

Proof. Let L be a positive or negative modal predicate logic, and $\mathfrak{M}$ its canonical model. Assume some set Γ of $\mathfrak{L}$-formulas is L-consistent. By the extensibility lemma 9.5, Γ is contained in some Henkin set w for L. (Note that Γ contains no variables from Var^*.) By the truth lemma 9.7, $\mathfrak{M}, w, g_w \Vdash A$ for each $A \in \Gamma$. So Γ is satisfiable any class of structures that contains the structure of $\mathfrak{M}$.

Lemma 9.9 (Canonicity of FK^*).
The structure of the canonical model for FK^* is total.

Proof. Immediate from lemma 9.4.

Theorem 9.10 (Completeness of FK^*).
The system FK^* is strongly complete with respect to the class of total counterpart structures.

Proof. Immediate from lemmas 9.8 and 9.9.

Lemma 9.11 (Canonicity of NK^*).
The structure of the canonical model for NK is single-domain.

Proof. Immediate from lemmas 9.4.

> **Theorem 9.12 (Completeness of** NK*).
> *The system* NK* *is strongly complete with respect to the class of single-domain counterpart structures.*

Proof. Immediate from lemmas 9.8 and 9.11.

Let me briefly return to a point I mentioned on p.13: that the introduction of multiple counterpart relations makes little difference to the base logics. The easiest way to see this is perhaps to note that all the lemmas in the previous section still go through if we define accessibility and counterparthood in canonical models by a fixed transformation τ whose range excludes infinitely many variables. The extensibility lemma (9.5) and existence lemma (9.6) are unaffected by this change; the only part that needs adjusting is the clause for $\Box B$ in the proof of the truth lemma 9.7, but the adjustments are straightforward.

Bibliography

Tom Baldwin [1984]: "Lowe on Modalities de Re". *Mind*, 93(370): 252–255

Francesco Belardinelli [2006]: *Quantified Modal Logic and the Ontology of Physical Objects*. PhD Thesis

— [2022]: "Counterpart Semantics at Work: Independence and Incompleteness Results in Quantified Modal Logic". In Francesco Ademollo, Fabrizio Amerini and Vincenzo De Risi (Eds.) *Thinking and Calculating,* Cham: Springer, 409–431

Patrick Blackburn, Maarten de Rijke and Yde Venema [2001]: *Modal Logic*. Cambridge: Cambridge University Press

Patrick Blackburn and Johan van Benthem [2007]: "Modal Logic: A Semantic Perspective". In *Handbook of Modal Logic,* Elsevier, 1–84

Torben Braüner and Silvio Ghilardi [2007]: "First-Order Modal Logic". In *Handbook of Modal Logic,* 549–620

Rudolf Carnap [1947]: *Meaning and Necessity*. Chicago: University of Chicago Press

Giovanna Corsi [2002a]: "Counterpart Semantics. A Foundational Study on Quantified Modal Logics"

— [2002b]: "A Unified Completeness Theorem for Quantified Modal Logics". *The Journal of Symbolic Logic*, 67(4): 1483–1510

— [2007]: "Necessary For". In *Proceedings of the LMPS,* Bejing

Max J. Cresswell [1995]: "Incompleteness and the Barcan Formula". *Journal of Philosophical Logic*, 24: 379–403

John Divers [2007]: "Quinean Scepticism About De Re Modality After David Lewis". *European Journal of Philosophy*, 15(1): 40–62

Melvin Fitting and Richard L. Mendelsohn [1998]: *First-Order Modal Logic*. Dordrecht: Kluwer

James W. Garson [2001]: "Quantification in Modal Logic". Synthese Library. Dordrecht: Springer Netherlands, 249–307

Olivier Gasquet [1995]: "A New Incompleteness Result in Kripke Semantics". *Fundamenta Informaticae*, 24(4): 407–415

Silvio Ghilardi [1989]: "Presheaf Semantics and Independence Results for Some Non-Classical First-Order Logics". *Archive for Mathematical Logic*, 29(2): 125–136

— [1991]: "Incompleteness Results in Kripke Semantics". *Journal of Symbolic Logic*, 56(2): 517–538

— [2001]: "Substitution, Quantifiers and Identity in Modal Logic". In E. Morscher and A. Hieke (Eds.) *New Essays in Free Logic,* Dordrecht: Kluwer, 87–115

Silvio Ghilardi and Giancarlo Meloni [1988]: "Modal and Tense Predicate Logic: Models in Presheaves and Categorical Conceptualization". In *Categorical Algebra and Its Applications,* 130–142

— [1991]: "Philosophical and Mathematical Investigations in First-Order Modal Logic". In *Problemi Fondazionali in Teoria Del Significato. Atti Del Convegno Di Pontignano,* Firenze: Olsckhi, 77–107

Allan Gibbard [1975]: "Contingent Identity". *Journal of Philosophical Logic*, 4: 187–221

Katherine Hawley [2001]: *How Things Persist*. Oxford University Press

Allen P. Hazen [1977]: *The Foundations of Modal Logic*. PhD Thesis, University of Pittsburgh

— [1979]: "Counterpart-Theoretic Semantics for Modal Logic". *Journal of Philosophy*, 76: 319–338

Michael Townsen Hicks [2022]: "Counterparts and Counterpossibles: Impossibility without Impossible Worlds". *The Journal of Philosophy*, 119(10): 542–574

G. E. Hughes and Max J. Cresswell [1996]: *A New Introduction to Modal Logic*. London and New York: Routledge

Graeme Hunter and William Seager [1981]: "The Discreet Charm of Counterpart Theory". *Analysis*, 41(2): 73–76

Toomas Karmo [1983]: "Contingent Non-Identity". *Australasian Journal of Philosophy*, 61(2): 185–187

Alexander W. Kocurek [2018]: "Counteridenticals". *The Philosophical Review*, 127(3): 323–369

Markus Kracht and Oliver Kutz [2002]: "The Semantics of Modal Predicate Logic I. Counterpart Frames". In Frank Wolter, Heinrich Wansing, Maarten de Rijke and Michael Zakharayaschev (Eds.) *Advances in Modal Logic,* vol 3. World Scientific Publishing Company

— [2005]: "The Semantics of Modal Predicate Logic II. Modal Individuals Revisited". In Reinhard Kahle (Ed.) *Intensionality,* Los Angeles: A. K. Peters

— [2007]: "Logically Possible Worlds and Counterpart Semantics for Modal Logic". In Dale Jacquette (Ed.) *Philosophy of Logic,* Amsterdam: North Holland, 943–995

Saul Kripke [1963]: "Semantical Considerations on Modal Logics". *Acta Philosophica Fennica*, 16: 83–94

— [1971]: "Identity and Necessity". In Milton K. Munitz (Ed.) *Identity and Individuation,* New York: New York University Press, 135–164

Steven T Kuhn [1980]: "Quantifiers as Modal Operators". *Studia Logica. An International Journal for Symbolic Logic*, 39(2-3): 145–158

Oliver Kutz [2000]: "Kripke-Typ-Semantiken Für Die Modale Prädikatenlogik"

Karel Lambert [2017]: "Free Logics". In *The Blackwell Guide to Philosophical Logic,* chapter 12. John Wiley & Sons, Ltd, 258–279

Bibliography

William F. Lawvere [1969]: "Adjointness in Foundations". *Dialectica*, 23: 281–296

David Lewis [1971]: "Counterparts of Persons and Their Bodies". *Journal of Philosophy*, 68: 203–211

— [1973]: *Counterfactuals*. Oxford: Blackwell

— [1983]: "Postscript to "Counterpart Theory and Quantified Modal Logic"". In *Philosophical Papers I,* Oxford: Oxford University Press, 39–46

— [1986]: *On the Plurality of Worlds*. Malden (Mass.): Blackwell

David K. Lewis [1968]: "Counterpart Theory and Quantified Modal Logic:". *Journal of Philosophy*, 65(5): 113–126

Neil McDonnell [2016]: "Events and Their Counterparts". *Philosophical Studies*, 173(5): 1291–1308

Dilip Ninan [2018]: "Quantification and Epistemic Modality". *The Philosophical Review*, 127(4): 433–485

John Nolt [2021]: "Free Logic". In Edward N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy,* Metaphysics Research Lab, Stanford University, fall 2021 edition

Hiroakira Ono [1973]: "Incompleteness of Semantics for Intermediate Predicate Logics. I Kripke's Semantics". *Proceedings of the Japan Academy*, 49(9): 711–713

Murali Ramachandran [2020]: "Chisholm's Modal Paradox(Es) and Counterpart Theory 50 Years On". *Logic and Logical Philosophy*, 29(4): 571–592

Gerhard Schurz [2011]: "Combinations and Completeness Transfer for Modal Quantificational Logics". *Logic Journal of the IGPL*, 19(4): 598–616

Wolfgang Schwarz [2014]: "Counterpart Theory and the Paradox of Occasional Identity". *Mind*, 123(492): 1057–1094

Hiroyuki Shirasu [1998]: "Duality in Superintuitionistic and Modal Predicate Logics"

Dmitrij Skvortsov and Valentin B. Shehtman [1993]: "Maximal Kripke-type Semantics for Modal and Superintuitionistic Predicate Logics". *Annals of Pure and Applied Logic*, 63(1): 69–101

Robert Stalnaker and Richmond Thomason [1968]: "Abstraction in First Order Modal Logic". *Theoria*, 34

Lin Woollaston [1994]: "Counterpart Theory as a Semantics for Modal Logic". *Logique et Analyse*, 147–148: 255–263