WEAKLY CLASSICAL THEORIES OF IDENTITY

JOSHUA B. SCHECHTER

Abstract. There are well-known quasi-formal arguments that identity is a "strict" relation in at least the following three senses: (1) There is a single identity relation and a single distinctness relation; (2) There are no contingent cases of identity or distinctness; and (3) There are no vague or indeterminate cases of identity or distinctness. However, the situation is less clear cut than it at first may appear. There is a natural formal theory of identity that is very close to the standard classical theory but which does not validate the formal analogues of (1)–(3). The core idea is simple: We weaken the Principle of the Indiscernibility of Identicals from a conditional to an entailment and we adopt a weakly classical logic. This paper investigates this weakly classical theory of identity (and related theories) and discusses its philosophical ramifications. It argues that we can accept a reasonable theory of identity without committing ourselves to the uniqueness, necessity, or determinacy of identity.

§1. Introduction. The identity relation seems to be very straightforward. If asked to characterize this relation, we might respond with a near triviality – for example, "everything is identical to itself and to nothing else".[1] No more elaborate characterization appears needed.

The formal theory of identity also seems to be well-settled. The familiar first-order theory of identity is generated by two principles – the claim that identity is reflexive, and a schematic version of Leibniz's Principle of the Indiscernibility of Identicals.

Part of the reason that the nature of identity seems so straightforward is that identity appears to be a "strict" relation. In particular, identity seems to be strict in at least the following three senses:

(1) There is a single identity relation and a single distinctness (non-identity) relation. There is no room for multiple extensionally-inequivalent identity or distinctness relations.

(2) The identity and distinctness relations obtain necessarily. There is no room for contingent cases of identity or distinctness.

---

[1]For instance, Lewis writes, "Identity is utterly simple and unproblematic. Everything is identical to itself; nothing is ever identical to anything else except itself. There is never any problem about what makes something identical to itself; nothing can ever fail to be. And there is never any problem about what makes two things identical; two things never can be identical." See Lewis (1986), pages 192–193.

(3) The identity and distinctness relations obtain definitely and determinately. There is no room for vague or indeterminate cases of identity or distinctness.[2]

Indeed, there are compelling quasi-formal arguments for each of these three claims.

To be sure, there can be indeterminate and contingent identity *statements*. Well-known examples include the contingent "Benjamin Franklin is identical to the first Postmaster General"[3] and the indeterminate "Princeton is identical to Princeton Borough" (as opposed to the surrounding Princeton Township).[4] But in these cases, the contingency or indeterminacy seems not to derive from the identity relation as such. Rather, the contingency and indeterminacy seems to derive from the ways in which terms in the two sentences designate their referents. The description "the first Postmaster General" is not a rigid designator. It does not designate the same object – namely, Benjamin Franklin – in every possible world. The term "Princeton" does not determinately refer to a single geographic region. It refers indeterminately to Princeton Borough, Princeton Township, and perhaps other geographical regions. When we restrict our attention to singular terms that designate their objects both rigidly and determinately, examples of contingent or indeterminate identity statement are hard to find. According to orthodoxy, then, (1)–(3) are true.

However, the technical situation is less clear cut than it may at first appear. There is a natural theory of identity that provides more wiggle room than one might have thought possible. This theory is very close to the standard classical theory of identity, but does not validate the formal analogues of (1)–(3). The core idea is simple: We weaken the Principle of the Indiscernibility of Identicals from a conditional to an entailment (or equivalently, to a rule). We also adopt a weakly classical logic – a logic that validates the classical entailments but not the classical meta-rules. In the resulting theory, the standard arguments for the formal analogues of (1)–(3) do not go through. Weaker versions of the formal results do obtain. But they do not rule out the possibility of multiple extensionally-inequivalent identity relations or the possibility of contingent or indeterminate cases of identity.

Perhaps this result will turn out to be a mere logical curiosity. But it provides evidence that we can adopt a reasonable theory of identity without claiming that identity is unique, necessary, or determinate. If we have reason – from metaphysics, perhaps – to reject one of these claims, we can do so without radically changing our logic or our theory of identity. At the very least, we can stay neutral on these issues.

Indeed, I suspect that this result tells us something deep and interesting about the nature of identity and the power of classical logic. It shows us where the strength of the familiar theory of identity comes from. And this is so even if we ultimately reject weakly classical theories as insufficiently powerful.

---

[2]In this paper, I treat vagueness as a species of indeterminacy. So far as I can tell, nothing hangs on this assumption.

[3]See Kripke (1971).

[4]See Lewis (1988).

This paper will proceed as follows: In the next section, I rehearse the quasi-formal arguments for the strictness of identity. In section three, I develop the proof theory and model theory for weakly classical theories of identity. I first discuss a simple theory of identity that contains the claim that identity is reflexive and a version of the Principle of the Indiscernibility of Identicals in rule form. I then consider theories that contain additional principles. I provide a natural algebraic semantics for each of these theories and prove strong soundness and strong completeness results. In section four, I examine how to add a modal operator to the language. In section five, I present an alternative relational semantics. In section six, I return to the quasi-formal arguments for the strictness of identity and show how they break down. Finally, in section seven, I briefly discuss the potential philosophical ramifications of these results.

§2. **Arguments for the Strictness of Identity.** Here are the basic principles of the familiar classical theory of identity:[5]

$x = x$ Reflexivity (Re)

$x = y \to (E(x) \leftrightarrow E(y))$ Indiscernibility of Identicals^c (PII^c)

These two principles are schemas such that x and y are arbitrary terms and E is an arbitrary formula. For simplicity, we can assume that the language lacks function symbols. Thus, the only terms are constants and variables. The formula $E(y)$ differs from the formula $E(x)$ only in that it may have a free occurrence of y in some or all of the places that $E(x)$ has a free occurrence of x. (We treat all occurrences of constants as free.)

2.1. The Collapse Argument. The core of the collapse argument is a simple formal result. Suppose we have a language with two identity predicates, $=_1$ and $=_2$. Suppose they each obey the classical theory of identity. That is, suppose that the theory contains:

$x =_1 x$ Reflexivity$_1$ (Re$_1$)

$x =_2 x$ Reflexivity$_2$ (Re$_2$)

$x =_1 y \to (E(x) \leftrightarrow E(y))$ Indiscernibility of Identicals$^c_1$ (PII$^c_1$)

$x =_2 y \to (E(x) \leftrightarrow E(y))$ Indiscernibility of Identicals$^c_2$ (PII$^c_2$)

Here, $E(x)$ ranges over formulas of the entire language.

Given this theory, we can show that $=_1$ and $=_2$ are fully intersubstitutable. That is, any pair of formulas that differ only in the subscripts on some or all of their identity symbols logically entail each other.

PROOF. Let x and y be arbitrary constants or variables. By Re$_2$, $\vdash x =_2 x$. By PII$^c_1$, $\vdash x =_1 y \to (x =_2 x \leftrightarrow x =_2 y)$. Since $x =_2 x, x =_1 y \to (x =_2 x \leftrightarrow x =_2 y) \vdash x =_1 y \to x =_2 y$, it follows that $\vdash x =_1 y \to x =_2 y$. Similarly, using Re$_1$ and PII$^c_2$, $\vdash x =_2 y \to x =_1 y$. So, $\vdash x =_1 y \leftrightarrow x =_2 y$.

In classical logic, we have the following principle: If $\vdash A \leftrightarrow B$ then $\vdash E \leftrightarrow E\langle B/A \rangle$, where A, B, and E are arbitrary formulas and $E\langle B/A \rangle$ differs from E only in that it has occurrences of B in some or all of the places where E has occurrences of A. (This is the principle of the substitution of logical equivalents.)

---

[5] In what follows, I follow the standard practice of ignoring the use-mention distinction where there is little chance of confusion.

So for any formula E, $\vdash E \leftrightarrow E\langle x =_2 y / x =_1 y\rangle$. It follows that $E \dashv\vdash E\langle x =_2 y / x =_1 y\rangle$. Thus, $=_1$ and $=_2$ are fully intersubstitutable.[6] $\square$

This result can be used to argue for several striking claims. First, it can be used to argue that two people who each speak a language with an identity predicate must have identity predicates with the very same extension.[7] In particular, if the speakers disagree about the truth of an identity claim, at least one of them must be wrong. For otherwise, we could combine their languages. The two identity predicates would presumably each obey the classical theory of identity in the combined language but would be inequivalent. This violates the formal result.[8]

Second, the formal result can be used to argue that if a speaker employs an identity predicate, it determinately refers to a single determinate relation. For otherwise, there would be at least two inequivalent precisifications of the speaker's predicate. These precisifications could be added to the speaker's language. The precisifications would presumably each obey the classical theory of identity in the expanded language but would be inequivalent. This again violates the formal result.

Finally, and perhaps more surprisingly, the formal result can be used to argue that identity obtains necessarily (for any reasonable kind of alethic necessity). For otherwise, we could define a second identity predicate, $='$, such that $x =' y$ holds just in case $\Box x = y$, where $\Box$ is the relevant necessitation operator. Since $\vdash x = x$, we have it that $\vdash \Box x = x$ (by the Necessitation principle of modal logic).[9] Moreover, $\vdash \Box x = y \rightarrow x = y$ (by the T principle of modal logic). Using PII^c, $\vdash \Box x = y \rightarrow (E(x) \leftrightarrow E(y))$. So $='$ obeys the classical theory of identity but is inequivalent with $=$. This again violates the formal result.

One might worry that these arguments prove too much. After all, there are cases of contingent and indeterminate identity statements. What should we say about them? The answer is that the Principle of the Indiscernibility of Identicals should not be taken to apply in contexts where designators can fail to be rigid or fail to determinately refer. In what follows, we'll assume that this constraint is met – all terms in the language will be assumed to designate both rigidly and determinately.

Each of the three ways of applying the collapse argument can be challenged. For example, against the first argument, it might be claimed that one cannot always include the identity predicates of distinct languages in a common language. Alternatively, it might be claimed that when a language with an identity predicate is expanded with additional vocabulary, there may come to be a violation of the Principle of the Indiscernibility of Identicals where there wasn't one originally. Against the second argument, it might be claimed that there are cases

---

[6] A version of this argument can be found in Williamson (2007). He traces it back to Quine (1961).

[7] If distinct predicates can have the same extension but different meanings, the two identity predicates may nevertheless have different meanings.

[8] An analogous argument can be made for mental representations rather than public language expressions.

[9] One might object to this on the grounds that $\Box a = a$ is equivalent to $\Box \exists x x = a$, and so $\Box a = a$ entails that a necessarily exists. Following Kripke (1971), we can avoid this worry by taking $\Box$ to be a weak necessitation operator: $\Box A$ is true if A is true in every world in which the objects mentioned in A exist.

of indeterminacy without distinct precisifications. Alternatively, it might be claimed that precisifications of the identity predicate can violate the Principle of the Indiscernibility of Identicals with respect to formulas containing occurrences of distinct precisifications. Against the third argument, it might be claimed that the Principle of the Indiscernibility of Identicals fails in modal contexts. And so on. As is probably obvious, these challenges carry a whiff of desperation about them, at least in the absence of significant defense and elaboration.

2.2. The Modal Argument. The second quasi-formal argument for the strictness of identity consists of a cluster of related results in classical identity theory with a single one-place modal operator. If the modal operator, $\Box$, is governed by a normal modal logic, we can show that $\vdash x = y \to \Box x = y$.

PROOF. By Re, $\vdash x = x$. By Necessitation, $\vdash \Box x = x$. By PII^c, $\vdash x = y \to (\Box x = x \to \Box x = y)$. Since $\Box x = x, x = y \to (\Box x = x \to \Box x = y) \vdash x = y \to \Box x = y$, we have it that $\vdash x = y \to \Box x = y$.[10] $\qquad \Box$

This proof relies only on the Necessitation principle of modal logic. Given additional modal principles, we can derive further results. For example, consider the B principle: $\vdash A \to \Box\neg\Box\neg A$. In the modal logic KB, we can show that $\vdash \neg x = y \to \Box\neg x = y$.

PROOF. By the previous result, $\vdash x = y \to \Box x = y$. Since $\vdash x = y \leftrightarrow \neg\neg x = y$, using K it follows that $\vdash \Box x = y \leftrightarrow \Box\neg\neg x = y$. So $\vdash x = y \to \Box\neg\neg x = y$. By contraposition, $\vdash \neg\Box\neg\neg x = y \to \neg x = y$. Again using K, $\vdash \Box\neg\Box\neg\neg x = y \to \Box\neg x = y$. By B, $\vdash \neg x = y \to \Box\neg\Box\neg\neg x = y$. Therefore, $\vdash \neg x = y \to \Box\neg x = y$. $\qquad \Box$

Putting these two results together, it follows that in any normal extension of KB, the necessity of identity and the necessity of distinctness obtain. It further follows that $\vdash \neg\Box x = y \to \Box\neg x = y$ and $\vdash \neg\Box\neg x = y \to \Box x = y$. That is, if an identity is not necessary, its negation is necessary. If the negation of an identity is not necessary, the identity is necessary. Moreover, using the Law of the Excluded Middle, $\vdash \Box x = y \lor \Box\neg x = y$. In other words, for any identity claim, either it or its negation is a necessary truth.

Now consider the D principle: $\vdash \Box A \to \neg\Box\neg A$. In the modal logic KD, we can show that $\vdash \Box\neg x = y \to \neg x = y$.

PROOF. By D, $\vdash \Box x = y \to \neg\Box\neg x = y$. By contraposition, $\vdash \Box\neg x = y \to \neg\Box x = y$. By the above results, $\vdash x = y \to \Box x = y$. By contraposition, $\vdash \neg\Box x = y \to \neg x = y$. Therefore, $\vdash \Box\neg x = y \to \neg x = y$. $\qquad \Box$

In the modal logic KDB, we can show that $\vdash \Box x = y \to x = y$.

PROOF. By D, $\vdash \Box x = y \to \neg\Box\neg x = y$. By the above results, in any extension of KB, $\vdash \neg x = y \to \Box\neg x = y$. By contraposition, $\vdash \neg\Box\neg x = y \to x = y$. Therefore, $\vdash \Box x = y \to x = y$. $\qquad \Box$

---

[10]This argument is commonly referred to as the "Barcan-Kripke argument". The first proof of the necessity of identity appears in Barcan (1947). A version of this argument is also stated in Kripke (1971). Kripke doesn't explicitly endorse the argument (although he does endorse its conclusion). Kripke writes that the argument "has been stated many times in recent philosophy."

Thus, in any normal extension of KDB, $x = y$ and $\Box x = y$ are logically equivalent, as are $\neg x = y$ and $\Box \neg x = y$. In particular, these results apply to the modal logic S5 – the best candidate for the logic of metaphysical necessity. This suggests that we should accept the necessity of identity and the necessity of distinctness. There is no room for any kind of contingency in the identity and distinctness relations.

2.3. The Vagueness Argument. The final argument for the strictness of identity that I'll consider is a version of Evans's argument that there can be no vague objects.[11]

Let V be an operator such that $\mathrm{V}\,A$ stands for the claim that it is vague whether A. We first show that $\vdash \mathrm{V}\,x = y \to \neg x = y$. The only principle for V that we'll need is that it is not vague whether logical truths obtain.

PROOF. By Re, $\vdash x = x$. By the non-vagueness of logical truths, $\vdash \neg \mathrm{V}\,x = x$. By PII^c, $\vdash x = y \to (\neg \mathrm{V}\,x = x \leftrightarrow \neg \mathrm{V}\,x = y)$. Since $\neg \mathrm{V}\,x = x, x = y \to (\neg \mathrm{V}\,x = x \leftrightarrow \neg \mathrm{V}\,x = y) \vdash x = y \to \neg \mathrm{V}\,x = y$, it follows that $\vdash x = y \to \neg \mathrm{V}\,x = y$.[12] By contraposition, $\vdash \mathrm{V}\,x = y \to \neg x = y$. $\qquad\square$

This suggests that if it is vague whether an identity holds, the identity is false. Presumably, if an identity is false it is not vague whether it holds.[13] Thus, we can conclude that identity cannot be vague.

If it is vague whether $\neg x = y$, it is vague whether $x = y$. So we should endorse $\vdash \mathrm{V}\,\neg x = y \to \mathrm{V}\,x = y$. Combining this with the above result, $\vdash \mathrm{V}\,\neg x = y \to \neg x = y$. This suggests that if it is vague whether a non-identity holds, the identity is false. Presumably, if an identity is false it is not vague whether the non-identity holds.[14] Thus, we can conclude that non-identity cannot be vague, either.

There is a somewhat different (and perhaps better) strategy that can be used to argue that identity and distinctness are not vague.[15] In particular, we can apply the results of the previous subsection. Suppose Det is an operator such that $\mathrm{Det}\,A$ stands for the claim that it is definitely the case that A. (Thus, we can understand $\mathrm{V}\,A$ to be equivalent to $\neg\,\mathrm{Det}\,A \wedge \neg\,\mathrm{Det}\,\neg A$.) If Det obeys the Necessitation principle of modal logic (as seems plausible), then by the results of the previous subsection, $\vdash x = y \to \mathrm{Det}\,x = y$. If, in addition, Det obeys the principles of the modal logic KB (as also seems plausible), then $\vdash \neg x = y \to \mathrm{Det}\,\neg x = y$. In KB, then, both identities and non-identities are definite. From this, it follows that $\vdash \neg\,\mathrm{Det}\,x = y \to \mathrm{Det}\,\neg x = y$ and $\vdash \neg\,\mathrm{Det}\,\neg x = y \to \mathrm{Det}\,x = y$. That is, if it is not definite that an identity holds, it is definite that it does

---

[11]See Evans (1978). See Salmon (1981), page 243, for essentially the same argument. Evans's main concern is whether the world can contain vague objects and not whether identity can be vague. As Williamson (1994), page 256, points out, it is plausible that we can accept the existence of vague objects without accepting vague identities. So even if Evans's argument works against vague identity, it may not rule out the existence of vague objects.

[12]Notice that the argument to this point is analogous to the Barcan-Kripke argument, with $\neg \mathrm{V}$ playing the role of $\Box$. The analogy is not perfect, however, since $\neg \mathrm{V}\,A$ is not the claim that A is definitely true but the claim that it not vague whether A is true.

[13]This claim can be questioned.

[14]This claim can also be questioned.

[15]This may be related to what Evans had in mind with the second argument in his paper.

not hold. If it is not definite that an identity does not hold, it is definite that the identity does hold. Moreover, by the Law of the Excluded Middle, it follows that $\vdash \mathrm{Det}\, x = y \vee \mathrm{Det}\, \neg x = y$. In other words, for any identity statement, either it or its negation is definite. (And therefore, both $\vdash \neg\, \mathrm{V}\, x = y$ and $\vdash \neg\, \mathrm{V}\, \neg x = y$.) This suggests that we should accept the definiteness of identity and distinctness. There is no room for vagueness in the identity or distinctness relations.[16]

Similarly, if we interpret $\mathrm{V}\, A$ as the claim that it is indeterminate whether A and $\mathrm{Det}\, A$ as the claim that it is determinately the case that A, we can argue that there is no room for any kind of indeterminacy in the identity or distinctness relations.

2.4. Weakly Classical Identity Theory. What should we make of these arguments? One option is to simply accept their conclusions. This may ultimately be the best response. But it seems much too quick to close off our options in metaphysics on the basis of these rather superficial-seeming grounds. It is worth investigating how the arguments may be resisted.

A second option is to move to a very different base logic. There are several non-classical logics that could be explored. For instance, we could move to an intuitionistic logic.[17] If we did that, we would still be able to prove that $\vdash x =_1 y \leftrightarrow x =_2 y$, that $\vdash x = y \rightarrow \Box x = y$, and that $\vdash x = y \rightarrow \mathrm{Det}\, x = y$. But we would no longer be able to prove that $\vdash \Box x = y \vee \Box \neg x = y$ or that $\vdash \mathrm{Det}\, x = y \vee \mathrm{Det}\, \neg x = y$. Alternatively, we could move to a logic in which $B, A \rightarrow (B \rightarrow C) \nvdash A \rightarrow C$.[18] In such a logic, the proofs of $\vdash x =_1 y \leftrightarrow x =_2 y$, $\vdash x = y \rightarrow \Box x = y$, and $\vdash x = y \rightarrow \mathrm{Det}\, x = y$ would fail. Needless to say, these are radical proposals.

A third option is to retain classical logic but reject the standard first-order theory of identity. The trouble with this suggestion is that it is not at all clear what should go in its place.

The strategy of this paper is to try to make the *minimal* possible changes in logic and identity theory and still escape the arguments for the strictness of identity. It turns out that two small changes suffice.

The first change is to weaken the theory of identity. Reflexivity seems non-negotiable. The idea, then, is to slightly weaken the Principle of the Indiscernibility of Identicals. The natural suggestion is to weaken the principle from a conditional to an entailment (or to a rule). There are two salient proposals for doing so:

$$x = y, E(x)/E(y) \qquad \mathrm{PII}^e$$
$$x = y/E(x) \leftrightarrow E(y) \qquad \mathrm{PII}^{e\prime}$$

These two rules are equivalent over a very weak base logic. In particular, all that is needed beyond the usual structural rules are two principles for the biconditional: $A, A \leftrightarrow B \vdash B$ and $\vdash A \leftrightarrow A$.

---

[16]Indeed, it is plausible that Det obeys the principles of the modal logic KDB. By the results of the previous subsection, it follows that $x = y$ and $\mathrm{Det}\, x = y$ are logically equivalent, as are $\neg x = y$ and $\mathrm{Det}\, \neg x = y$.

[17]See Putnam (1983) and Wright (2001) for the suggestion that the correct account of vagueness requires moving to an intuitionistic logic.

[18]Such logics are sometimes proposed to deal with the semantic paradoxes. See, for example, Field (2008).

PROOF. By PII^e, $x = y, E(x) \leftrightarrow E(x) \vdash E(x) \leftrightarrow E(y)$. But $\vdash E(x) \leftrightarrow E(x)$. So $x = y \vdash E(x) \leftrightarrow E(y)$.

By PII$^{e\prime}$, $x = y \vdash E(x) \leftrightarrow E(y)$. $E(x), E(x) \leftrightarrow E(y) \vdash E(y)$. So $x = y, E(x) \vdash E(y)$. $\qquad\square$

We can therefore adopt either of these two rules as our version of the Principle of the Indiscernibility of Identicals. For concreteness, we'll adopt PII^e.[19] This is a familiar rule – it is the standard elimination rule for identity in Natural Deduction-style formulations of identity theory.

In (strongly) classical logic, PII^c and PII^e are equivalent. The second change we need to make is to slightly weaken our base logic. In particular, we weaken the logic to weakly classical logic.

Given a consequence relation for a language that contains at least the usual propositional logical constants, we say that the consequence relation is *weakly classical* just in case each classically valid formula is valid and each classically valid entailment is valid. (For the purpose of this definition, we do not count identity as a logical constant. So PII^c does not count as a classical validity.) We say that the consequence relation is *strongly classical* if, in addition, each classically valid meta-rule is valid. In particular, the consequence relation must have as valid Conditional Introduction, Disjunction Elimination (that is, Reasoning by Cases), and Negation Introduction (that is, Reductio).

Over weakly classical logic, Conditional Introduction, Disjunction Elimination, and Negation Introduction are equivalent. These meta-rules are also equivalent to the principle of Entailment Congruence: If $\Gamma, A \vdash B$ and $\Gamma, B \vdash A$ then $\Gamma, E \vdash E\langle B/A\rangle$. So we can characterize a strongly classical logic as a weakly classical logic that validates any one of these meta-rules.

If we're working in a language in which there are only the usual logical connectives and quantifiers (in addition to individual constants, functions, and predicates), any weakly classical consequence relation is strongly classical. But if the language contains additional vocabulary, a weakly classical consequence relation need not be strongly classical. For instance, there are weakly classical theories of truth and weakly classical theories of determinateness that are not strongly classical.[20] Perhaps surprisingly, the same is true for identity. As we will see below, if our theory of identity is axiomatized by Re and PII^e, the weakly classical consequence relation will not be strongly classical.[21] This provides just enough wiggle room to block the arguments for (1)–(3).

§3. Basic Results.

3.1. The Language. We can begin our investigation in a very simple language. The language has an infinite stock of individual constants, $a_1, a_2, a_3, \ldots$, the familiar logical connectives, $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$, the identity predicate, $=$, and

---

[19]The superscript "c" stands for "conditional" and the superscript "e" stands for "entailment".

[20]Logics for truth or determinate truth based on supervaluational semantics provide important examples of this.

[21]By contrast, by Lemma 3.15, if our theory of identity consists of Re and PII^c, the consequence relation will be strongly classical.

the parentheses, (and). The language does not contain variables, quantifiers, function symbols, or any additional predicates.

We use a, b, and c to stand for arbitrary constants. We use A, B, and C to stand for arbitrary sentences. We use Γ, Δ, and Θ to stand for arbitrary sets of sentences.

The formation rules for sentences are the obvious ones. The only atomic sentences are expressions of the form $a = b$, where a and b are constants. If A and B are sentences, so are $\neg A$, $(A \wedge B)$, $(A \vee B)$, $(A \to B)$, and $(A \leftrightarrow B)$. Nothing else is a sentence.

3.2. Proof Theory. We operate with a weakly classical consequence relation. So we can make use of any of the standard Hilbert-style axiomatizations of classical propositional logic for our language.[22] For simplicity, we can assume that the only rule of inference in the chosen axiomatization is Modus Ponens. In addition, we add the following two schemas:

$a = a$ Reflexivity (Re)

$a = b, E(a)/E(b)$ Indiscernibility of Identicals^e (PII^e)

The sentence $E(b)$ differs from the sentence $E(a)$ only in that it may have occurrences of b in some or all of the places that $E(a)$ has occurrences of a.

We let $\top$ stand for an arbitrary theorem of the logic.

3.3. Semantics. We employ an algebraic approach to semantics. Each model for the language is based on a Boolean algebra. There are two ways of understanding what these Boolean algebras represent. On one way of understanding the semantics, the elements of the Boolean algebras are truth values. On an alternative – and preferable – way of understanding the semantics, the elements of the Boolean algebras are propositions. They are not truth values but the contents expressed by sentences.

There is a choice to make regarding the designated values of our semantics. We could insist that the sole designated value in a model be the top element of the relevant Boolean algebra. Alternatively, we could make use of a proper filter of elements of the Boolean algebra – a non-empty proper upper subset of the algebra that is closed under finite infima.

If we understand the elements of the Boolean algebra to be a set of propositions, we can understand this choice as follows: On one view, a proposition is a set of possible worlds, precisifications, or some other kind of entity. Employing a semantics where the designated value is the top element is akin to a supervaluational semantics on which a sentence is true if it is true on all precisifications. Employing a semantics where the designated values form a proper filter of propositions is akin to a modal semantics on which a sentence is true if it expresses a set of worlds that contains the actual world (or some privileged set of worlds).[23] In what follows, I'll discuss both approaches, but I'll largely

---

[22]We could instead make use of a Natural Deduction-style formulation of classical propositional logic, so long as the meta-rules were restricted so that PII^e cannot be applied in subderivations.

[23]More precisely, there is a natural correspondence between (i) a complete atomic Boolean algebra with a complete proper filter of designated values and (ii) a set of worlds with a privileged subset of worlds to supervaluate over. If the filter contains only the top element of the Boolean algebra, the corresponding privileged set of worlds is the set of all worlds. If the

focus on the filter-based approach. This is for two reasons. First, it is philo-
sophically more natural in most contexts. Second, there are technical reasons for
making use of a filter-based semantics when we extend the language by adding
an alethic modal operator. In particular, on the most straightforward way of
adding a modal operator to the single designated value semantics, it turns out
that if A is true in a model so is $\Box A$. When we turn to the case of vagueness
and indeterminacy, however, the single designated value approach will become
relevant.

In greater detail, here is the filter-based semantics: A model, $\mathcal{M}$, is an ordered
sextuple $\langle S, \leq, F, D, I, \Phi \rangle$, such that $\langle S, \leq \rangle$ is a non-trivial[24] Boolean algebra. F
is a proper filter of the Boolean algebra. D is a non-empty set – the domain of
the model. I is a function from $D \times D$ to S – the identity function of the model.
Φ is a function from the set of constants to D – the interpretation function of
the model.

We write p and q for arbitrary elements of S. We write d, e, and f for arbitrary
elements of the domain.

Given any sentence A of the language, we write $\|A\|^{\mathcal{M}}$ for the semantic value
of A in the model $\mathcal{M}$. When the model is clear from context, we omit the
superscript. The semantic value of A in $\mathcal{M}$ is recursively defined as follows:

- $\|a = b\| = I(\Phi(a), \Phi(b))$;
- $\|\neg A\| = -\|A\|$;
- $\|A \wedge B\| = \|A\| \sqcap \|B\|$;
- $\|A \vee B\| = \|A\| \sqcup \|B\|$;
- $\|A \to B\| = -\|A\| \sqcup \|B\|$;
- $\|A \leftrightarrow B\| = (-\|A\| \sqcup \|B\|) \sqcap (-\|B\| \sqcup \|A\|)$.

Here, $-$, $\sqcap$, and $\sqcup$ are the complement, infimum, and supremum functions for
the Boolean algebra, respectively. We write 0 and 1 for the bottom and top
elements of the algebra, respectively.

Given a model $\mathcal{M}$, we write $\mathcal{M} \vDash A$ to mean that A is true in $\mathcal{M}$. This obtains
just in case $\|A\| \in F$. We write $\mathcal{M} \vDash \Gamma$ to mean that every member of Γ is true
in $\mathcal{M}$.

Notice that in any model, $\|A \to B\| = 1$ just in case $\|A\| \leq \|B\|$. $\|A \leftrightarrow B\| = 1$
just in case $\|A\| = \|B\|$.

In this semantics, we'll want theorems to have value 1 in each model (and
not merely a designated value). Otherwise, we can find a model in which the
theorem is not designated – we simply take a model in which a theorem has a
value other than 1 and modify it by setting the filter to contain only 1.

We must impose constraints on models so that our theory of identity comes
out as true in every model. In particular, there are two constraints that are
needed:

- $I(d, d) = 1$;
- If $I(d, e) \in F$ then for any f, $I(d, f) = I(e, f)$ and $I(f, d) = I(f, e)$.

---

filter is an ultrafilter, the corresponding privileged set of worlds is a singleton. See section 5.3
for relevant discussion.

[24]That is, S has at least two elements.

The first constraint is needed so that each instance of Reflexivity has value 1 in every model. The second constraint is needed for PII^e. In what follows, models will be understood to satisfy these two constraints. (We'll add further constraints below.)

Notice that the two constraints jointly entail that if $I(d,e) \in F$ then $I(d,e) = 1$: Suppose $I(d,e) \in F$. Then $I(d,e) = I(e,e)$ (by the second constraint) $= 1$ (by the first constraint). Thus, in any model, a designated identity statement has the top value in the Boolean algebra.

We write $\vDash A$ to mean that A is valid. This obtains just in case for every model $\mathcal{M}$, $\mathcal{M} \vDash A$. We write $\Gamma \vDash A$ to mean that Γ entails A. This obtains just in case for every model $\mathcal{M}$, if $\mathcal{M} \vDash \Gamma$ then $\mathcal{M} \vDash A$.[25]

The single designated value semantics is similar. We simply add the constraint that in any model, the only element of F is the top element of its Boolean algebra.

3.4. Soundness. It is straightforward to show that the proof theory is strongly sound with respect to our filter-based semantics. That is, if $\Gamma \vdash A$ then $\Gamma \vDash A$.

LEMMA 3.1. *In any model, if* $\|a = b\| \in F$ *then* $\|E(a)\| = \|E(b)\|$.

PROOF. Suppose $\|a = b\| \in F$. We prove the result by induction on the complexity of E.

- $E(a)$ is $a = c$: If $E(b)$ is $a = c$, the result is trivial. Suppose instead that $E(b)$ is $b = c$. $\|E(a)\| = \|a = c\| = I(\Phi(a), \Phi(c)) = I(\Phi(b), \Phi(c))$ (by the second constraint) $= \|b = c\| = E(b)$.
- The cases where $E(a)$ is $c = a$ or $a = a$ are similar.
- $E(a)$ is $\neg G(a)$. $\|E(a)\| = \|\neg G(a)\| = -\|G(a)\| = -\|G(b)\|$ (by induction) $= \|\neg G(b)\| = \|E(b)\|$.
- $E(a)$ is $G(a) \wedge H(a)$. $\|E(a)\| = \|G(a) \wedge H(a)\| = \|G(a)\| \sqcap \|H(a)\| = \|G(b)\| \sqcap \|H(b)\|$ (by induction) $= \|G(b) \wedge H(b)\| = \|E(b)\|$.
- The remaining cases are similar. □

LEMMA 3.2. *If* $\vdash A$ *then in any model,* $\|A\| = 1$.

PROOF. Suppose $\vdash A$. Let $\mathcal{M}$ be any model. We show that $\|A\| = 1$ by induction on the complexity of the proof of A.

- Propositional Logic. If A is a theorem of classical propositional logic, then $\|A\| = 1$, since the model is based on a Boolean algebra.
- Reflexivity: $\|a = a\| = I(\Phi(a), \Phi(a)) = 1$ (by the first constraint).
- PII^e: Suppose $\|a = b\| = 1$ and $\|E(a)\| = 1$. By Lemma 3.1, $\|E(b)\| = \|E(a)\| = 1$.
- Modus Ponens: Suppose $\|A \to B\| = 1$. So $\|A\| \leq \|B\|$. So if $\|A\| = 1$ then $\|B\| = 1$. □

THEOREM 3.3 (Strong Soundness). *If* $\Gamma \vdash A$ *then* $\Gamma \vDash A$.

PROOF. Suppose $\Gamma \vdash A$. Let $\mathcal{M}$ be any model such that $\mathcal{M} \vDash \Gamma$. We show that $\mathcal{M} \vDash A$ by induction on the derivation of A from Γ.

- $A \in \Gamma$: Since $\mathcal{M}$ is a model of Γ, $\mathcal{M} \vDash A$.
- A is an axiom: By Lemma 3.2, $\mathcal{M} \vDash A$.

---

[25]Entailment is thus defined to be global rather than local validity.

- Modus Ponens: Suppose A is derived from B and $B \to A$ via an application of Modus Ponens. By induction, $\mathcal{M} \vDash B$ and $\mathcal{M} \vDash B \to A$. So $\|B\| \in F$ and $\|B \to A\| \in F$. Since F is a filter, $\|B\| \sqcap \|B \to A\| \in F$. $\|B\| \sqcap \|B \to A\| = \|B\| \sqcap (-\|B\| \sqcup \|A\|) = (\|B\| \sqcap -\|B\|) \sqcup (\|B\| \sqcap \|A\|) = 0 \sqcup (\|B\| \sqcap \|A\|) = \|B\| \sqcap \|A\| \leq \|A\|$. Since F is a filter, $\|A\| \in F$.
- PII^e: Suppose A is $E(b)$ and is derived from $a = b$ and $E(a)$ via an application of PII^e. By induction, $\mathcal{M} \vDash a = b$. By Lemma 3.1, $\|E(a)\| = \|E(b)\|$. By induction, $\mathcal{M} \vDash E(a)$. So $\mathcal{M} \vDash E(b)$. □

We are now in a position to show that PII^c is not valid on the semantics.

PROPOSITION 3.4. *PII^c is not valid.*

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m, n\}$; $I(l, m) = I(m, l) = I(l, n) = I(n, l) = s$ and $I(m, n) = I(n, m) = t$; and $\Phi(a_1) = l$, $\Phi(a_2) = m$, and $\Phi(a_3) = n$.

In this model, $\|a_1 = a_2 \to (a_1 = a_3 \leftrightarrow a_2 = a_3)\| = -s \sqcup (t \sqcap s) = t \notin F$. □

On the semantics, PII^c does not follow from PII^e because $a = b, E(a) \vdash E(b)$ "says" that if $a = b$ and $E(a)$ get designated values, so does $E(b)$. The equivalent principle $\text{PII}^{e\prime}$ "says" that if $a = b$ gets a designated value, so does $E(a) \leftrightarrow E(b)$. But this tells us nothing about what happens if $a = b$ gets an undesignated value. By contrast, PII^c requires $E(a) \leftrightarrow E(b)$ to have at least as great a value as $a = b$, whether or not $a = b$ is designated.

Since PII^e is valid but PII^c is not, it immediately follows that Conditional Introduction is not a valid meta-rule. Thus:

COROLLARY 3.5. *The consequence relation $\vDash$ is weakly classical but not strongly classical.*

3.5. Completeness. We can prove that the proof theory is strongly complete with respect to this algebraic semantics using a familiar kind of Lindenbaum-Tarski construction.

Given a set of sentences of the language, Δ, we define the following relation on pairs of sentences in the language: $A \sim^\Delta B$ just in case $\Delta \vdash A \leftrightarrow B$.

LEMMA 3.6. *$\sim^\Delta$ is an equivalence relation on sentences.*

PROOF. Easy. □

Let $|A|^\Delta$ be the equivalence class of the sentence A generated by $\sim^\Delta$. We make the following definitions:

$\mathbf{S}^\Delta = \{|A|^\Delta \mid A \text{ is a sentence}\}$;

$|A|^\Delta \leq^\Delta |B|^\Delta$ just in case $\Delta \vdash A \to B$.

Our proof of completeness relies on the following well-known result:[26]

LEMMA 3.7. *$\langle \mathbf{S}^\Delta, \leq^\Delta \rangle$ is a Boolean algebra in which $-|A|^\Delta = |\neg A|^\Delta$, $|A|^\Delta \sqcap |B|^\Delta = |A \wedge B|^\Delta$, $|A|^\Delta \sqcup |B|^\Delta = |A \vee B|^\Delta$, the top element $\mathbf{1}^\Delta = |\top|^\Delta$, and the bottom element $\mathbf{0}^\Delta = |\neg\top|^\Delta$. $\langle \mathbf{S}^\Delta, \leq^\Delta \rangle$ is non-trivial just in case Δ is consistent.*

---

[26]See, for instance, Rasiowa & Sikorski (1970), chapter 6.

Given any set of sentences, Γ, we define the following relation on pairs of terms in the language: $a \equiv^\Gamma b$ just in case $\Gamma \vdash a = b$.

LEMMA 3.8. $\equiv^\Gamma$ *is an equivalence relation on terms.*

PROOF. By Re, $\vdash a = a$. So $\Gamma \vdash a = a$. So $\equiv^\Gamma$ is reflexive.

Suppose $\Gamma \vdash a = b$. By PII^e, $a = b, a = a \vdash b = a$. By R, $\vdash a = a$. So $a = b \vdash b = a$. So $\Gamma \vdash b = a$. Therefore, $\equiv^\Gamma$ is symmetric.

Suppose $\Gamma \vdash a = b$ and $\Gamma \vdash b = c$. By the above, $a = b \vdash b = a$. By PII^e, $b = a, b = c \vdash a = c$. So $\Gamma \vdash a = c$. Therefore, $\equiv^\Gamma$ is transitive. $\qquad\square$

For any constant a, let a^Γ be the equivalence class containing a generated by $\equiv^\Gamma$. Let $\Gamma^=$ be the set of identity sentences derivable from Γ. Given any set of sentences, Γ, and any set Δ such that $\Gamma^= \subseteq \Delta \subseteq \Gamma$, we make the following definitions:

$\mathbf{F}^{\Gamma,\Delta} = \{|A|^\Delta \mid \Gamma \vdash A\}$;

$\mathbf{D}^\Gamma = \{a^\Gamma \mid a$ is a constant in the language$\}$;

$\mathbf{I}^{\Gamma,\Delta}(a^\Gamma, b^\Gamma) = |a = b|^\Delta$;

$\mathbf{\Phi}^\Gamma(a) = a^\Gamma$;

$\mathbf{M}^{\Gamma,\Delta} = \langle \mathbf{S}^\Delta, \leq^\Delta, \mathbf{F}^{\Gamma,\Delta}, \mathbf{D}^\Gamma, \mathbf{I}^{\Gamma,\Delta}, \mathbf{\Phi}^\Gamma \rangle$.

We can use $\mathbf{M}^{\Gamma,\Delta}$ as our model of Γ in the proof of strong completeness.

LEMMA 3.9. *If Γ is consistent, $\mathbf{M}^{\Gamma,\Delta}$ is a model.*

PROOF. $\langle \mathbf{S}^\Delta, \leq^\Delta \rangle$ is a non-trivial Boolean algebra: Since Γ is consistent, so is Δ. By Lemma 3.7, $\langle \mathbf{S}^\Delta, \leq^\Delta \rangle$ is a non-trivial Boolean algebra.

$\mathbf{F}^{\Gamma,\Delta}$ is well-defined: If $A \sim^\Delta B$ then $\Delta \vdash A \leftrightarrow B$. So $\Gamma \vdash A \leftrightarrow B$. So $\Gamma \vdash A$ just in case $\Gamma \vdash B$.

$\mathbf{F}^{\Gamma,\Delta}$ is non-empty: $\Gamma \vdash \top$. So $\mathbf{1}^\Delta = |\top|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$.

$\mathbf{F}^{\Gamma,\Delta}$ is an upper subset of the Boolean algebra: Suppose $|A|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$ and $|A|^\Delta \leq^\Delta |B|^\Delta$. Then $\Gamma \vdash A$ and $\Delta \vdash A \rightarrow B$. Since $\Delta \subseteq \Gamma$, $\Gamma \vdash A \rightarrow B$. So $\Gamma \vdash B$. So $|B|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$.

$\mathbf{F}^{\Gamma,\Delta}$ is a filter: Suppose $|A|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$ and $|B|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$. So $\Gamma \vdash A$ and $\Gamma \vdash B$. So $\Gamma \vdash A \wedge B$. So $|A \wedge B|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$. By Lemma 3.7, $|A|^\Delta \sqcap |B|^\Delta = |A \wedge B|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$. So $\mathbf{F}^{\Gamma,\Delta}$ is closed under finite infima.

$\mathbf{F}^{\Gamma,\Delta}$ is a proper filter: Since Γ is consistent, $\Gamma \nvdash A$ for some A. So $|A|^\Delta \notin \mathbf{F}^{\Gamma,\Delta}$.

Clearly, $\mathbf{D}^\Gamma$ is a non-empty set and $\mathbf{\Phi}^\Gamma$ is a function from the constants to $\mathbf{D}^\Gamma$.

$\mathbf{I}^{\Gamma,\Delta}$ is well-defined: Suppose $a^\Gamma = c^\Gamma$. So $\Gamma \vdash a = c$. So $\Gamma^= \vdash a = c$. Since $\Gamma^= \subseteq \Delta$, $\Delta \vdash a = c$. By PII$^{e\prime}$, $a = c \vdash a = b \leftrightarrow c = b$. So $\Delta \vdash a = b \leftrightarrow c = b$. So $|a = b|^\Delta = |c = b|^\Delta$. Similarly, if $b^\Gamma = c^\Gamma$ then $|a = b|^\Delta = |a = c|^\Delta$.

Finally, we must show that $\mathbf{M}^\Gamma$ satisfies the two constraints on models.

Constraint 1: $\vdash a = a$. So $\vdash a = a \leftrightarrow \top$. So $\Delta \vdash a = a \leftrightarrow \top$. Therefore, $\mathbf{I}^{\Gamma,\Delta}(a^\Gamma, a^\Gamma) = |a = a|^\Delta = |\top|^\Delta = \mathbf{1}^\Delta$.

Constraint 2: Suppose $\mathbf{I}^{\Gamma,\Delta}(a^\Gamma, b^\Gamma) \in \mathbf{F}^{\Gamma,\Delta}$. So $|a = b|^\Delta \in \mathbf{F}^{\Gamma,\Delta}$. So $\Gamma \vdash a = b$. So $a^\Gamma = b^\Gamma$. Therefore, $\mathbf{I}^{\Gamma,\Delta}(a^\Gamma, c^\Gamma) = \mathbf{I}^{\Gamma,\Delta}(b^\Gamma, c^\Gamma)$ and $\mathbf{I}^{\Gamma,\Delta}(c^\Gamma, a^\Gamma) = \mathbf{I}^{\Gamma,\Delta}(c^\Gamma, b^\Gamma)$. $\qquad\square$

LEMMA 3.10. *If Γ is consistent, $\mathbf{M}^{\Gamma,\Delta} \vDash A$ just in case $\Gamma \vDash A$.*

PROOF. Suppose Γ is consistent. By Lemma 3.9, $\mathbf{M}^{\Gamma,\Delta}$ is a model. We show that $\|A\| = |A|^{\Delta}$ by induction on the complexity of A.

- $\|a = b\| = \mathbf{I}^{\Gamma,\Delta}(\mathbf{\Phi}^{\Gamma}(a), \mathbf{\Phi}^{\Gamma}(b)) = \mathbf{I}^{\Gamma,\Delta}(a^{\Gamma}, b^{\Gamma}) = |a = b|^{\Delta}$.
- $\|\neg A\| = -\|A\| = -|A|^{\Delta}$ (by induction) $= |\neg A|^{\Delta}$.
- $\|A \wedge B\| = \|A\| \sqcap \|B\| = |A|^{\Delta} \sqcap |B|^{\Delta}$ (by induction) $= |A \wedge B|^{\Delta}$.
- The cases of $\vee$, $\rightarrow$, and $\leftrightarrow$ are similar.

Therefore, $\mathbf{M}^{\Gamma,\Delta} \vDash A$ just in case $\|A\| \in \mathbf{F}^{\Gamma,\Delta}$ just in case $|A|^{\Delta} \in \mathbf{F}^{\Gamma,\Delta}$ just in case $\Gamma \vdash A$. $\square$

Given these lemmas, we can show the following:

THEOREM 3.11 (Strong Completeness). *If $\Gamma \vDash A$ then $\Gamma \vdash A$.*

PROOF. Suppose $\Gamma \nvdash A$. So Γ is consistent. By Lemma 3.9, $\mathbf{M}^{\Gamma,\Delta}$ is a model. By Lemma 3.10, $\mathbf{M}^{\Gamma,\Delta} \vDash \Gamma$ and $\mathbf{M}^{\Gamma,\Delta} \nvDash A$. Therefore, $\Gamma \nvDash A$. $\square$

COROLLARY 3.12 (Compactness). *If $\Gamma \vDash A$ then for some finite $\Gamma_0 \subseteq \Gamma$, $\Gamma_0 \vDash A$.*

PROOF. Suppose $\Gamma \vDash A$. By strong completeness, $\Gamma \vdash A$. So for some finite $\Gamma_0 \subseteq \Gamma$, $\Gamma_0 \vdash A$. By strong soundness, $\Gamma_0 \vDash A$. $\square$

COROLLARY 3.13. *Γ is consistent just in case Γ is satisfiable.*

PROOF. If Γ is consistent, then by Lemma 3.9, $\mathbf{M}^{\Gamma,\Delta}$ is a model. By Lemma 3.10, $\mathbf{M}^{\Gamma,\Delta}$ satisfies Γ.

If Γ is satisfiable, then $\Gamma \nvDash A$ for some sentence A. By strong soundness, $\Gamma \nvdash A$. So Γ is consistent. $\square$

For concreteness, we can take $\mathbf{M}^{\Gamma,\Gamma^{=}}$ to be our official Lindenbaum-Tarski model for Γ for the filter-based semantics.

Notice that the strong completeness result still obtains if we add the constraint that the only designated value in a model is the top value of the Boolean algebra. This follows from the fact that the model $\mathbf{M}^{\Gamma,\Gamma}$ has only a single designated value.

PROOF. Suppose $|A|^{\Gamma} \in \mathbf{F}^{\Gamma,\Gamma}$ and $|B|^{\Gamma} \in \mathbf{F}^{\Gamma,\Gamma}$. So $\Gamma \vdash A$ and $\Gamma \vdash B$. $A, B \vdash A \leftrightarrow B$. So $\Gamma \vdash A \leftrightarrow B$. Therefore $|A|^{\Gamma} = |B|^{\Gamma}$. $\square$

We can therefore take $\mathbf{M}^{\Gamma,\Gamma}$ to be our official Lindenbaum-Tarski model for Γ for the single designated value semantics.

3.6. Comments. On the algebraic semantics, the identity relation is highly constrained when an identity statement is designated. In any model, all designated identity statements take the top value of the Boolean algebra. This is the residual sense in which identity is a strict relation in the semantics. However, there is no constraint when an identity statement is undesignated. Identity statements can take arbitrary undesignated semantic values. This is the significant wiggle room that is allowed by the semantics.

We can add a further requirement to the definition of a model without changing the consequence relation generated by the semantics. In particular, we can require that $I(d, e) \in F$ just in case d is identical to e. (Here, "is identical to" is the identity relation in the meta-language.) This is a very natural requirement

to impose. Depending on one's view of the role of model theory, it might be thought to be a non-negotiable constraint.

Notice that the construction used in our proof of strong completeness, $\mathbf{M}^{\Gamma,\Delta}$, satisfies this requirement. (That was the reason for setting the domain of the constructed model to contain equivalence classes of constants.)

PROOF. $\mathbf{I}^{\Gamma,\Delta}(a^\Gamma, b^\Gamma) \in \mathbf{F}^{\Gamma,\Delta}$ just in case $\Gamma \vdash a = b$ just in case $a^\Gamma = b^\Gamma$. □

So the proof of strong completeness (and related results) goes through even given this additional requirement.

If we impose this requirement, we can replace the two constraints on models listed above with a single constraint:

- $I(d, e) \in F$ just in case $I(d, e) = 1$ just in case d is identical to e.

In what follows, we'll impose this constraint on models.

There are other natural constraints that could also be imposed. For instance, we could require that the Boolean algebra of each model be complete – that is, closed under arbitrary infima and suprema – and atomic – that is, such that every non-zero element is greater than an atom, a non-zero element immediately greater than zero. If we were to impose these constraints, we'd presumably also want to require that the filter be complete – that is, closed under arbitrary infima. Adding these constraints would not change the consequence relation. This is a consequence of the results stated in section 5.

3.7. Symmetry and Transitivity. There is a problem with the consequence relation that we have so far characterized. There are intuitively correct claims about identity that turn out not to be valid. For instance, identity is intuitively symmetric. It is intuitively transitive. However, $\nvdash a = b \rightarrow b = a$ and $\nvdash (a = b \land b = c) \rightarrow a = c$.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m, n\}$; $I(l, m) = I(m, n) = I(n, m) = s$, $I(l, n) = I(n, l) = I(m, l) = t$; and $\Phi(a_1) = l$, $\Phi(a_2) = m$, and $\Phi(a_3) = n$.

In this model, $\|a_1 = a_2 \rightarrow a_2 = a_1\| = -s \sqcup t = t \sqcup t = t \notin F$. Moreover, $\|(a_1 = a_2 \land a_2 = a_3) \rightarrow a_1 = a_3\| = -(s \sqcap s) \sqcup t = -s \sqcup t = t \notin F$. □

Weaker versions of symmetry and transitivity do hold. In particular, $a = b \vdash b = a$. $a = b, b = c \vdash a = c$. $a = b \vdash b = c \rightarrow a = c$. But the stronger principles are intuitively correct, too. This is a significant flaw with the consequence relation.

The obvious fix is to add additional principles to the theory of identity. The new theory of identity can be axiomatized using four basic principles:

$a = a$	Reflexivity (Re)
$a = b \rightarrow b = a$	Symmetry (Sy)
$(a = b \land b = c) \rightarrow a = c$	Transitivity (Tr)
$a = b, E(a)/E(b)$	Indiscernibility of Identicals^e (PII^e)

If we add the Sy and Tr principles, we need to impose additional constraints in the definition of a model. In total, we impose three constraints on models:

- $I(d, e) \in F$ just in case $I(d, e) = 1$ just in case d is identical to e;
- $I(d, e) = I(e, d)$;

- $I(d, e) \sqcap I(e, f) \leq I(d, f)$.

The second constraint is intuitively correct. The third constraint is also intuitively correct, since higher in the Boolean algebra corresponds, in some sense, to "truer".[27] The third constraint says, in effect, that $I(d, f)$ is at least as true as $I(d, e) \sqcap I(e, f)$.

It may be helpful to think about these three constraints as follows: Suppose we look not at identity but at distinctness. Let $\mathrm{NI}(d, e)$ be $-I(d, e)$. Then the constraints are equivalent to the following:

- $\mathrm{NI}(d, e) = 0$ just in case d is identical to e;
- $\mathrm{NI}(d, e) = \mathrm{NI}(e, d)$;
- $\mathrm{NI}(d, f) \leq \mathrm{NI}(d, e) \sqcup \mathrm{NI}(e, f)$;
- $I(d, e) \in F$ just in case $I(d, e) = 1$.

The first three of these constraints say that NI behaves like a metric, except that the range of the function is a Boolean algebra rather than the non-negative real numbers. (The third constraint is just the triangle inequality.) The constraints on models are thus very natural.

It is easy to show that the new consequence relation is strongly sound with respect to the algebraic semantics with the new constraints imposed on models. The second constraint ensures that Sy is valid and the third constraint ensures that Tr is valid. The strong completeness proof is also straightforward. The Lindenbaum-Tarski construction goes just as before. It is easy to show that the model $\mathbf{M}^{\Gamma, \Delta}$ satisfies the new constraints. Sy ensures that the model satisfies the second constraint and Tr ensures that it satisfies the third.

The resulting logic, however, has a problematic feature. It is a strongly classical logic.

LEMMA 3.14. *The minimal weakly classical consequence relation for our language that obeys Re, Sy, Tr, and PII^e also obeys PII^c.*

PROOF. We prove this by induction on the complexity of $E(a)$. The base cases are the cases where $E(a)$ is $a = a$, $a = c$, or $c = a$ for some distinct constant c. Each of these cases is easy. For illustration, here is one of them:

- $E(a)$ is $a = c$: The case where $E(b)$ is $a = c$ is trivial. Suppose instead that $E(b)$ is $b = c$. By Tr, $\vdash (a = b \land b = c) \to a = c$. So $\vdash a = b \to (b = c \to a = c)$. That is, $\vdash a = b \to (E(b) \to E(a))$. Again by Tr, $\vdash (c = a \land a = b) \to c = b$. So $\vdash a = b \to (c = a \to c = b)$. By Sy, $\vdash a = c \leftrightarrow c = a$ and $\vdash b = c \leftrightarrow c = b$. So, $\vdash a = b \to (a = c \to b = c)$. That is, $\vdash a = b \to (E(a) \to E(b))$. Therefore, $\vdash a = b \to (E(a) \leftrightarrow E(b))$.

For the inductive step, suppose $\vdash a = b \to (E(a) \leftrightarrow E(b))$ and $\vdash a = b \to (G(a) \leftrightarrow G(b))$. It easily follows that $\vdash a = b \to (\neg E(a) \leftrightarrow \neg E(b))$, $\vdash a = b \to ((E(a) \land G(a)) \leftrightarrow (E(b) \land G(b)))$, $\vdash a = b \to ((E(a) \lor G(a)) \leftrightarrow (E(b) \lor G(b)))$, $\vdash a = b \to ((E(a) \to G(a)) \leftrightarrow (E(b) \to G(b)))$, and $\vdash a = b \to ((E(a) \leftrightarrow G(a)) \leftrightarrow (E(b) \leftrightarrow G(b)))$. Therefore, PII^c obtains.   □

---

[27]For instance, if elements of the Boolean algebra correspond to sets of possible worlds, the $\leq$ relation corresponds to the subset relation. So higher in the algebra corresponds to truth in a superset of worlds.

LEMMA 3.15. *Any weakly classical consequence relation with an axiomatization that has as its only rules Modus Ponens and PII^e is strongly classical just in case it obeys PII^c.*

PROOF. Suppose that a weakly classical consequence relation obeys PII^e but not PII^c. Then there is a counterexample to Conditional Introduction. So the consequence relation is not strongly classical.

Now suppose that a weakly classical consequence relation has an axiomatization that has as its only rules Modus Ponens and PII^e. Suppose it obeys PII^c. We show that if $\Gamma, A \vdash B$ then $\Gamma \vdash A \to B$. The argument is based on a simple modification of the usual inductive proof of the deduction theorem:

- $B \in \Gamma$: $\Gamma \vdash B$. $B \vdash A \to B$. So $\Gamma \vdash A \to B$.
- B is an axiom: $\vdash B$. $B \vdash A \to B$. So $\vdash A \to B$. So $\Gamma \vdash A \to B$.
- B is derived from C and $C \to B$ using Modus Ponens: By induction, $\Gamma \vdash A \to C$ and $\Gamma \vdash A \to (C \to B)$. $A \to C, A \to (C \to B) \vdash A \to B$. So $\Gamma \vdash A \to B$.
- B is derived from $a = b$ and $E(a)$ by PII^e: So $B = E(b)$. By induction, $\Gamma \vdash A \to a = b$ and $\Gamma \vdash A \to E(a)$. By PII^c, $\vdash a = b \to (E(a) \leftrightarrow E(b))$. $A \to a = b, A \to E(a), a = b \to (E(a) \leftrightarrow E(b)) \vdash A \to E(b)$. So $\Gamma \vdash A \to E(b)$.

Since the consequence relation is a weakly classical consequence relation that obeys Conditional Introduction, it is strongly classical. □

Given these two lemmas, the following result is immediate:

PROPOSITION 3.16. *The minimal weakly classical consequence relation for our language that obeys Re, Sy, Tr, and PII^e is strongly classical.*

In a sense, this is a nice result. It demonstrates that the weakly classical theory of identity that includes PII^e, Re, Sy, and Tr is not missing any logical truth concerning identity that can be stated in our language. On the other hand, the whole point of the endeavor was to avoid a strongly classical logic. We wanted to accept PII^e without also having to accept PII^c. So the result might seem to be very problematic.

This problem is not, however, very serious. The reason that we can prove PII^c is that the language we have been working with so far is expressively very limited. If we expand our language by adding a stock of predicates, for instance, the logic becomes a merely weakly classical logic. We will have it that $a = b \vdash Pa \leftrightarrow Pb$ but not $\vdash a = b \to (Pa \leftrightarrow Pb)$.

3.8. Adding Predicates. Let us add a stock of predicates to our language. In particular, for each natural number n, we add a (possibly empty) stock of n-place predicates, $P_1^n, P_2^n, \ldots$. In the expanded language, the atomic sentences are defined to include expressions of the form $a = b$ as well as expressions of the form $P_m^n a_1 \ldots a_n$, where P_m^n is an n-place predicate and $a_1, \ldots, a_n$ are constants. The formation rules for non-atomic sentences are as before.

The proof theory is as before. Our consequence relation is the minimal weakly classical consequence relation for the expanded language that includes PII^e, Re, Sy, and Tr. PII^e is now understood so that $E(a)$ ranges over all sentences in the expanded language. As before, this consequence relation can be axiomatized

using any Hilbert-style axiomatization of classical logic, the axioms schemas Re, Sy, and Tr, and the rule of inference PII^e. For simplicity, we assume that the only other rule of inference is Modus Ponens.

The semantics is also straightforward. As before, a model is an ordered sextuple $\langle S, \leq, F, D, I, \Phi \rangle$. The interpretation function, Φ, now maps constants to members of D and maps n-place predicates to functions from ordered n-tuples of members of D to S. The definition of the semantic value of a sentences in a model is as before with the addition of a clause for atomic sentences containing predicates:

- $\|P_m^n a_1 \ldots a_n\| = \Phi(P_m^n)(\Phi(a_1) \ldots \Phi(a_n))$.

The rest of the semantics, including the constraints on models, is just as before.[28]

The proof of strong soundness is a straightforward modification of the above proof. The proof of strong completeness also requires only straightforward modifications. In the definition of the model $\mathbf{M}^{\Gamma, \Delta}$, we now use as our interpretation function the function $\mathbf{\Phi}^{\Gamma, \Delta}$, defined as follows:

$\mathbf{\Phi}^{\Gamma, \Delta}(a) = a^{\Gamma}$;

$\mathbf{\Phi}^{\Gamma, \Delta}(P_m^n)(a_1^{\Gamma}, \ldots, a_n^{\Gamma}) = |P_m^n a_1 \ldots a_n|^{\Delta}$.

In the analogue of the proof of Lemma 3.9, we must show that $\mathbf{\Phi}^{\Gamma, \Delta}$ is well-defined. This is easy.

PROOF. Suppose $a_i^{\Gamma} = b^{\Gamma}$. So $\Gamma \vdash a_i = b$. So $\Gamma^= \vdash a_i = b$. So $\Delta \vdash a_i = b$. $a_i = b \vdash P_m^n a_1 \ldots a_i \ldots a_n \leftrightarrow P_m^n a_1 \ldots b \ldots a_n$. Therefore, $|P_m^n a_1 \ldots a_i \ldots a_n|^{\Delta} = |P_m^n a_1 \ldots b \ldots a_n|^{\Delta}$. □

In the analogue of the proof of Lemma 3.10, we must show that $\|P_m^n a_1 \ldots a_n\| = |P_m^n a_1 \ldots a_n|^{\Delta}$. This is easy, too.

PROOF. $\|P_m^n a_1 \ldots a_n\| = \mathbf{\Phi}^{\Gamma, \Delta}(P_m^n)(a_1^{\Gamma}, \ldots, a_n^{\Gamma}) = |P_m^n a_1 \ldots a_n|^{\Delta}$. □

The rest of the proof of strong completeness (and related results) goes through just as before.

If there is at least one predicate in the language, we can show that PII^c is not valid.

PROOF. For simplicity, we work in a language with a single 1-place predicate, P_1^1. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = s$; $\Phi(a_1) = l$, $\Phi(a_2) = m$, $\Phi(P_1^1)(l) = s$, $\Phi(P_1^1)(m) = t$.

In this model, $\|a_1 = a_2 \to (P_1^1 a_1 \leftrightarrow P_1^1 a_2)\| = -s \sqcup (t \sqcap s) = t \sqcup 0 = t \notin F$. □

So the consequence relation is weakly classical. Problem solved. We will therefore take this logic to be our base weakly classical logic with identity.

There is a remaining issue that deserves discussion. Consider the following schema: $E(a), \neg E(b) \vdash \neg a = b$. This is the contrapositive of PII^e. This schema is intuitively correct.[29] However, not every instance of this schema is valid in

---

[28]If we were to remove the requirement that $I(d, e) \in F$ just in case d is identical to e and return to the original two constraints, we would need to impose an additional constraint for predicates: If $I(d, e) \in F$ then $\Phi(P_m^n)(d_1, \ldots, d, \ldots, d_n) = \Phi(P_m^n)(d_1, \ldots, e, \ldots, d_n)$.

[29]Heck (1998), page 288, argues that if we accept PII^e, we should also accept its contrapositive.

our logic. (In general, in a weakly classical logic, while the contrapositive of a valid conditional is valid, the contrapositive of a valid entailment need not be valid.)

PROOF. We work in a language with a single 1-place predicate, P_1^1. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = s$; $\Phi(a_1) = d$, $\Phi(a_2) = e$, $\Phi(P_1^1)(d) = 1$, $\Phi(P_1^1)(e) = 0$.

In this model, $\|P_1^1 a_1\| = 1$, $\|\neg P_1^1 a_2\| = 1$, $\|\neg a_1 = a_2\| = t \notin F$. $\qquad\square$

We do get certain instances of the schema as valid in our logic. In particular, using Proposition 3.16, we can show that each instance of the schema where $E(a)$ is built out of constants, the identity relation, and the familiar logical connectives is valid. But we do not get such simple instances as $P_1^1 a, \neg P_1^1 b \vdash \neg a = b$.

To handle this problem, we can add another rule to the proof theory:

$\quad E(a), \neg E(b)/\neg a = b \qquad \text{PII}^{-1e}$

If we do this, we have to impose an additional constraint on models. The only obvious constraint to impose is the following: If $\|E(a)\| \in F$ and $-\|E(b)\| \in F$ then $-I(\Phi(a), \Phi(b)) \in F$. It is straightforward to prove strong soundness and strong completeness for the semantics with this additional constraint.

We can also show that PII^c is not valid on the new consequence relation.

PROOF. We work in a language with a single 1-place predicate, P_1^1. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, \Phi \rangle$, where $S = \{1, r, s, t, u, v, w, 0\}$ and $\leq$ is the relation depicted in the following Hasse diagram:
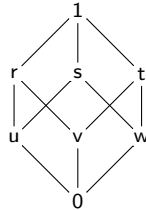
$F = \{1\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = r$; and $\Phi(a_1) = l$, $\Phi(a_2) = m$, $\Phi(P_1^1)(l) = r$, and $\Phi(P_1^1)(m) = s$.

This model satisfies the additional constraint: Suppose $\|E(a_1)\| = 1$. $E(a_1)$ is equivalent to a sentence in Conjunctive Normal Form – that is, a conjunction of disjunctions of atomic sentences and negations of atomic sentences. Each of the conjuncts has value 1. By a tedious enumeration of cases, it can be shown that substituting occurrences of a_2 for some of the occurrences of a_1 in a conjunct yields a sentence with value 1, r, or s. Thus, $E(a_2)$ has value 1, r, s, or u. So there is no pair of sentences $E(a_1)$ and $E(a_2)$ such that $E(a_1)$ and $\neg E(a_2)$ both have designated values. By a similar argument, there is no pair of sentences $E(a_1)$ and $E(a_2)$ such that $E(a_2)$ and $\neg E(a_1)$ both have designated values. So the model satisfies the constraint.

In this model, $\|a_1 = a_2 \to (P_1^1 a_1 \leftrightarrow P_1^1 a_2)\| = -r \sqcup (s \sqcap r) = w \sqcup u = s \notin F$. So PII^c is not valid. $\qquad\square$

The problem with adding PII^{-1e} to the logic is that the proposed additional constraint on models is unattractive. It concerns not only atomic sentences but also sentences of arbitrary complexity. As is clear from the proof immediately

above, verifying that a model satisfies the constraint can be tedious. And it is not clear that there is a simpler and more natural constraint that could instead be imposed. This is an unfortunate feature of the semantics.

In what follows, we will discuss both the base logic without PII^{-1e} and the logic that results from adding this rule.

As defined, our language lacks functions, variables, and quantifiers. If we add these to the language, our main results still obtain. The modifications needed in the definition of a model and in the proof of our main results are straightforward. They are the usual complications that arise for algebraic treatments of first-order logic.

§4. Adding a Modal Operator. Before we can return to the quasi-formal arguments for the strictness of identity, we must first examine how to add a modal operator to our language.

4.1. The Language. Our language is as before – containing constants and predicates but no functions, variables, or quantifiers – with the addition of a single one-place operator $\Box$. The atomic sentences are defined just as before. The non-atomic sentences are defined as before, with the addition of a clause for $\Box$: If A is a sentence, so is $\Box A$.

4.2. Proof Theory. The base logic for this language is the minimal weakly classical normal modal logic for our language that validates Re, Sy, Tr, and PII^e. A logic is a normal modal logic just in case the following obtains:

(K) $\vdash \Box(A \to B) \to (\Box A \to \Box B)$
(Nec)   If $\vdash A$ then $\vdash \Box A$

This base logic can therefore be axiomatized using any Hilbert-style axiomatization of classical logic, the axioms schemas Re, Sy, and Tr, an axiom schema for K, the rule of inference PII^e, and a rule of proof for Nec. (If we like, we can strengthen the logic by also adding the rule PII^{-1e}.)

Since we are working with an algebraic semantics, we could use a weaker modal base logic. For instance, we could adopt only the very weak modal principle RE: If $\vdash A \leftrightarrow B$ then $\vdash \Box A \leftrightarrow \Box B$. Alternatively, we could retain Nec and weaken K to K^e: $\Box(A \to B) \vdash \Box A \to \Box B$. We could weaken K still further to $\text{K}^{e'}$: $\Box(A \to B), \Box A \vdash \Box B$. For our purposes here, however, it will be convenient to work in a normal modal logic as our base logic.

There are several familiar modal principles that can be added to our base logic. For instance, we can add any of the following schemas:

(T) $\Box A \to A$
(4) $\Box A \to \Box\Box A$
(5) $\neg\Box A \to \Box\neg\Box A$
(B) $A \to \Box\neg\Box\neg A$
(D) $\Box A \to \neg\Box\neg A$

The most plausible modal logic for metaphysical necessity is S5 = KT45DB = K4DB = K5DB = KT4B = KT5. This logic is also plausibly a strengthening of the correct logic for definiteness (or determinacy). S5 will therefore be a particular focus of attention in what follows.

4.3. Semantics. The semantics for this language is straightforward. We enrich the definition of a model to handle sentences containing $\square$. The most natural way to do so is to add to each model a function N from S to S. The idea is that N maps the semantic value of a sentence to the semantic value that attributes necessity to the original semantic value.[30] A model, therefore, is an ordered septuple $\langle S, \leq, F, N, D, I, \Phi \rangle$, where $\langle S, \leq, F, D, I, \Phi \rangle$ is defined as above and N is a function from S to S.

The definition of the semantic value of a sentence A in a model is defined as above, with the addition of the obvious clause for $\square$:

- $\|\square A\| = N(\|A\|)$.

The definitions of truth in a model and validity are as before.

As before, we impose the following constraints on models:

- $I(d, e) \in F$ just in case $I(d, e) = 1$ just in case d is identical to e;
- $I(d, e) = I(e, d)$;
- $I(d, e) \sqcap I(e, f) \leq I(d, f)$.

If we add the rule PII^{-1e}, we also impose the following constraint:

- If $\|E(a)\| \in F$ and $-\|E(b)\| \in F$ then $-I(\Phi(a), \Phi(b)) \in F$.

We must impose additional constraints on models for the modal principles to be valid. It is easy to see that the semantics already validates RE. Indeed, it validates the more general principle:

(EQ)   If $\vdash A \leftrightarrow B$ then $\vdash E \leftrightarrow E\langle B/A \rangle$

However, additional constraints are required for Nec and K.

For Nec, we need to impose the following constraint on models: $N(1) = 1$. Notice that, given this constraint, if we require models to have only a single designated value then $A \vDash \square A$. This is not a plausible claim if $\square$ is interpreted to be an alethic modal operator. It is somewhat more plausible if $\square$ is interpreted to be a definiteness or determinately operator.

In discussing the case of K, it is helpful to consider a different modal principle:

(R) $\square(A \wedge B) \leftrightarrow (\square A \wedge \square B)$

In a weakly classical modal logic with Nec, K is equivalent to R.

PROOF. From K and Nec we can derive R: $\vdash (A \wedge B) \to A$. By Nec, $\vdash \square((A \wedge B) \to A)$. By K, $\vdash \square((A \wedge B) \to A) \to (\square(A \wedge B) \to \square A)$. By Modus Ponens, $\vdash \square(A \wedge B) \to \square A$. By analogous reasoning, $\vdash \square(A \wedge B) \to \square B$. It follows that $\vdash \square(A \wedge B) \to (\square A \wedge \square B)$.

$\vdash A \to (B \to (A \to B))$. By Nec, K, and Modus Ponens, $\vdash \square A \to \square(B \to (A \to B))$. By K, $\vdash \square(B \to (A \to B)) \to (\square B \to \square(A \to B))$. It follows that $\vdash \square A \to (\square B \to \square(A \wedge B))$ and so $\vdash (\square A \wedge \square B) \to \square(A \wedge B)$. Combining this with the above, $\vdash \square(A \wedge B) \leftrightarrow (\square A \wedge \square B)$.

From R and Nec we can derive K: $\vdash ((A \to B) \wedge A) \leftrightarrow (A \wedge B)$. By Nec and K, $\vdash \square((A \to B) \wedge A) \leftrightarrow \square(A \wedge B)$. Using R, $\vdash \square(A \wedge B) \to (\square A \wedge \square B)$. So $\vdash \square(A \wedge B) \to \square B$. So $\vdash \square((A \wedge B) \wedge A) \to \square B$. Using R, $\vdash (\square(A \to B) \wedge \square A) \to \square((A \to B) \wedge A)$. So $\vdash (\square(A \to B) \wedge \square A) \to \square B$. Therefore, $\vdash \square(A \to B) \to (\square A \to \square B)$. $\square$

---

[30] In algebraic semantics, it is more common to treat possibility rather than necessity as basic. But treating necessity as basic is slightly more convenient for our purposes here.

The obvious constraint corresponding to R is $N(p \sqcap q) = N(p) \sqcap N(q)$. For our base modal logic, we therefore impose two constraints on N in models:

- $N(1) = 1$;
- $N(p \sqcap q) = N(p) \sqcap N(q)$.

It is straightforward to determine the constraints on models that correspond to the other modal principles listed above.

- For T: $N(p) \leq p$;
- For 4: $N(p) \leq N(N(p))$;
- For 5: $-N(p) \leq N(-N(p))$;
- For B: $p \leq N(-N(-p))$;
- For D: $N(p) \leq -N(-p)$.

In particular, here is a natural list of constraints for S5:

- $N(1) = 1$;
- $N(p \sqcap q) = N(p) \sqcap N(q)$;
- $N(p) \leq p$;
- $N(p) \leq N(N(p))$;
- $-N(p) \leq N(-N(p))$.

Notice that the penultimate constraint is redundant.

There is an illuminating alternative characterization of the constraints for S5. Given a Boolean algebra, we say that the semantic value p is a Boolean combination of the semantic values $q_1, \ldots, q_n$ just in case p can be generated from the qs by taking complements, finite infima, and finite suprema. Given this definition, the last two constraints can be replaced with the following:

- If p is a Boolean combination of $N(p_1), \ldots, N(p_n)$ then $N(p) = p$.

PROOF. It is easy to see that this constraint entails the constraints corresponding to the 4 and 5 principles.

We show that the constraints listed for S5 entail this new constraint by induction on the complexity of the Boolean combination. It suffices to consider the base case and the cases of complement and pairwise infimum.

- $p = N(p_i)$: So $N(p) = N(N(p_i)) = N(p_i)$ (by the constraints for T and 4) $= p$.
- $p = -q$ where $q = N(q)$: $N(p) = N(-N(q)) = -N(q)$ (by the constraints for T and 5) $= q$.
- $p = q_1 \sqcap q_2$ where $q_1 = N(q_1)$ and $q_2 = N(q_2)$: $N(p) = N(q_1 \sqcap q_2) = N(q_1) \sqcap N(q_2)$ (by the constraint for R) $= q_1 \sqcap q_2 = p$.   □

Notice that the set of constraints for S5 is weaker than the constraint that $N(p) = 1$ if $p = 1$ and $N(p) = 0$ otherwise. This is not the appropriate constraint for S5 since any model that satisfies both this constraint and the constraint for PII^{-1e} obeys PII^c.[31]

PROOF. Suppose $\mathcal{M} \nvDash a = b \rightarrow (E(a) \leftrightarrow E(b))$. So $\|a = b \rightarrow (E(a) \leftrightarrow E(b))\| \notin F$. By the proposed constraint, $\|\Box(a = b \rightarrow (E(a) \leftrightarrow E(b)))\| = 0$. It follows that $-\|\Box(a = b \rightarrow (E(a) \leftrightarrow E(b)))\| = 1 \in F$. $\|\Box(a = a \rightarrow (E(a) \leftrightarrow E(a)))\| = 1 \in F$. By the constraint for PII^{-1e}, it follows that $-I(\Phi(a), \Phi(b)) \in F$. So

---

[31]See section 6.2 for the result that PII^c does not follow from S5 and PII^{-1e}.

$\|\neg a = b\| \in F$. Since $\|a = b \to (E(a) \leftrightarrow E(b))\| \notin F$, $\|\neg a = b\| \notin F$. This is a contradiction. $\square$

4.4. Soundness and Completeness. It is straightforward to prove the strong soundness of each of the weakly classical normal modal logics with identity described above with respect to our algebraic semantics with the corresponding constraints imposed on models.

It is also straightforward to prove the corresponding strong completeness results (and related results). We say that a set of sentences Δ is necessitation-closed just in case for every sentence A, if $\Delta \vdash A$ then $\Delta \vdash \Box A$. Given a set of sentences Γ and a necessitation-closed set of sentences Δ such that $\Gamma^= \subseteq \Delta \subseteq \Gamma$, we define the model $\mathbf{M}^{\Gamma,\Delta}$ to be $\langle \mathbf{S}^\Delta, \leq^\Delta, \mathbf{F}^{\Gamma,\Delta}, \mathbf{N}^\Delta, \mathbf{D}^\Gamma, \mathbf{\Phi}^\Gamma, \mathbf{I}^{\Gamma,\Delta} \rangle$, where $\langle \mathbf{S}^\Delta, \leq^\Delta, \mathbf{F}^{\Gamma,\Delta}, \mathbf{D}^\Gamma, \mathbf{\Phi}^\Gamma, \mathbf{I}^{\Gamma,\Delta} \rangle$ is defined as above and $\mathbf{N}^\Delta(|A|^\Delta) = |\Box A|^\Delta$. We use $\mathbf{M}^{\Gamma,\Delta}$ as our model for Γ in the proof of strong completeness.

We must make four additions to the proof of strong completeness. First, it must be shown that $\mathbf{N}^\Delta$ is well-defined. (This is the reason for requiring that Δ be necessitation-closed.)

PROOF. Suppose $|A|^\Delta = |B|^\Delta$. So $\Delta \vdash A \leftrightarrow B$. So $\Delta \vdash A \to B$ and $\Delta \vdash B \to A$. Since Δ is necessitation-closed, $\Delta \vdash \Box(A \to B)$ and $\Delta \vdash \Box(B \to A)$. By K, $\Delta \vdash \Box A \to \Box B$ and $\Delta \vdash \Box B \to \Box A$. Therefore, $\Delta \vdash \Box A \leftrightarrow \Box B$. $\square$

Second, it must be shown that for every Γ, there is a necessitation-closed Δ such that $\Gamma^= \subseteq \Delta \subseteq \Gamma$. This is easy: $\Gamma^=$ is itself a necessitation-closed set.

PROOF. Suppose $\Gamma^= \vdash A$. We show that $\Gamma^= \vdash \Box A$ by induction on the derivation of A.

- $A \in \Gamma^=$: A is $a = b$. By Re, $\vdash a = a$. By Nec, $\vdash \Box a = a$. By PII^e, $a = b, \Box a = a \vdash \Box a = b$. So $a = b \vdash \Box a = b$. So $\Gamma^= \vdash \Box a = b$.
- A is an axiom: $\vdash A$. By Nec, $\vdash \Box A$. So $\Gamma^= \vdash \Box A$.
- Modus Ponens: Suppose A is derived from B and $B \to A$ via an application of Modus Ponens. By induction, $\Gamma^= \vdash \Box B$ and $\Gamma^= \vdash \Box(B \to A)$. By K, $\Gamma^= \vdash \Box B \to \Box A$. So $\Gamma^= \vdash \Box A$.
- PII^e: Suppose A is $E(b)$ and is derived from $a = b$ and $E(a)$ via an application of PII^e. By induction, $\Gamma^= \vdash \Box E(a)$. By PII^e, $a = b, \Box E(a) \vdash \Box E(b)$. So $\Gamma^= \vdash \Box E(b)$. $\square$

Thus, we can use the model $\mathbf{M}^{\Gamma,\Gamma^=}$ as our model for Γ in the proof of completeness.

Third, it must be shown that $\mathbf{N}^\Delta$ satisfies the appropriate constraints for the given weakly classical modal logic with identity.

PROOF. Here is the proof for each of the modal principles listed above:

Nec: $\vdash \top$. By Nec, $\vdash \Box\top$. So $\vdash \Box\top \leftrightarrow \top$. So $\Delta \vdash \Box\top \leftrightarrow \top$. So $\mathbf{N}^\Delta(\mathbf{1}^\Delta) = \mathbf{N}^\Delta(|\top|^\Delta) = |\Box\top|^\Delta = |\top|^\Delta = \mathbf{1}^\Delta$.

R: By R, $\vdash \Box(A \wedge B) \leftrightarrow (\Box A \wedge \Box B)$. So $\Delta \vdash \Box(A \wedge B) \leftrightarrow (\Box A \wedge \Box B)$. So $\mathbf{N}^\Delta(|A|^\Delta \sqcap |B|^\Delta) = \mathbf{N}^\Delta(|A \wedge B|^\Delta) = |\Box(A \wedge B)|^\Delta = |\Box A \wedge \Box B|^\Delta = |\Box A|^\Delta \sqcap |\Box B|^\Delta = \mathbf{N}^\Delta(|A|^\Delta) \sqcap \mathbf{N}^\Delta(|B|^\Delta)$.

T: By T, $\vdash \Box A \to A$. So $\Delta \vdash \Box A \to A$. So $\mathbf{N}^\Delta(|A|)^\Delta = |\Box A|^\Delta \leq^\Delta |A|^\Delta$.

4: By 4, $\vdash \square A \rightarrow \square\square A$. So $\Delta \vdash \square A \rightarrow \square\square A$. So $\mathbf{N}^\Delta(|A|^\Delta) = |\square A|^\Delta \leq^\Delta$ $|\square\square A|^\Delta = \mathbf{N}^\Delta(\mathbf{N}^\Delta(|A|^\Delta))$.

5: By 5, $\vdash \neg\square A \rightarrow \square\neg\square A$. So $\Delta \vdash \neg\square A \rightarrow \square\neg\square A$. So $-\mathbf{N}^\Delta(|A|^\Delta) =$ $|\neg\square A|^\Delta \leq^\Delta |\square\neg\square A|^\Delta = \mathbf{N}^\Delta(-\mathbf{N}^\Delta(|A|^\Delta))$.

B: By B, $\vdash A \rightarrow \square\neg\square\neg A$. So $\Delta \vdash A \rightarrow \square\neg\square\neg A$. So $|A|^\Delta \leq^\Delta |\square\neg\square\neg A|^\Delta =$ $\mathbf{N}^\Delta(-\mathbf{N}^\Delta(-|A|^\Delta))$.

D: By D, $\vdash \square A \rightarrow \neg\square\neg A$. So $\Delta \vdash \square A \rightarrow \neg\square\neg A$. So $\mathbf{N}^\Delta(|A|^\Delta) = |\square A|^\Delta \leq^\Delta$ $|\neg\square\neg A|^\Delta = -\mathbf{N}^\Delta(-|A|^\Delta)$. $\qquad\square$

Finally, in the proof of the analogue of Lemma 3.10, it must be shown that $\|\square A\| = |\square A|^\Delta$. This is also easy.

PROOF. $\|\square A\| = \mathbf{N}^\Delta(\|A\|) = \mathbf{N}^\Delta(|A|^\Delta) = |\square A|^\Delta$. $\qquad\square$

The remainder of the strong completeness proof (and the proofs of related results) goes just as before.

If we add the rule $A/\square A$ to any of the weakly classical normal modal logics with identity described above, it will be strongly sound and strongly complete with respect to the semantics with the corresponding constraints on models and the additional constraint that the filter contain only the top value of the Boolean algebra. This is because any set Γ will be necessitation-closed given the addition of the new rule. So $\mathbf{M}^{\Gamma,\Gamma}$ is well-defined. This model has only a single designated value. So $\mathbf{M}^{\Gamma,\Gamma}$ can be used as the model for Γ in the proof of strong completeness for the single designated value semantics.

§5. **Relational Semantics.** We have so far made use of an algebraic approach to semantics. It is more common to make use of a relational approach, especially when discussing languages containing a modal operator. In this section, I present a relational semantics and state the constraints on models corresponding to each of the weakly classical theories of identity described above. I do so both because this style of semantics is more familiar and because it is illuminating to see what relational models look like for our weakly classical logics with identity.

In this section, I do not provide detailed proofs of strong soundness and strong completeness. I only present the main definitions and constructions needed for these (and related) results. The proofs are straightforward given the definitions and constructions.

5.1. Semantics. On the relational semantics, a model, $\mathcal{M}$, is an ordered sextuple $\langle W, R, @, D, I, \Phi \rangle$ where W is a non-empty set – the set of worlds. R is a binary relation on W – the accessibility relation. @ is a non-empty subset of W – the set of actual worlds.[32] D is a non-empty set – the domain of the model. I is a function that maps each world to a subset of $D \times D$ – the identity function of the model. Φ is a function that maps each constant to a member of

---

[32]This semantics thus bears some resemblance to the suggestion in Akiba (2004), Barnes (2009), and Williams (2008) that there are multiple actual worlds. Notice that we needn't think of W and @ as representing sets of worlds. They can be thought of representing sets of precisifications or of some other kind of entity.

D and maps each n-place predicate to a function that maps each world to a set of ordered n-tuples of members of D – the interpretation function of the model.

Given a model and a sentence A of the language we write $w \vDash A$ to mean that the sentence A is true at the world $w \in W$. This is recursively defined as follows:

- $w \vDash a = b$ just in case $\langle \Phi(a), \Phi(b) \rangle \in I(w)$;
- $w \vDash P_m^n a_1 \ldots a_n$ just in case $\langle \Phi(a_1), \ldots, \Phi(a_n) \rangle \in \Phi(P_m^n)(w)$;
- $w \vDash \neg A$ just in case $w \nvDash A$;
- $w \vDash A \wedge B$ just in case $w \vDash A$ and $w \vDash B$;
- $w \vDash A \vee B$ just in case $w \vDash A$ or $w \vDash B$;
- $w \vDash A \rightarrow B$ just in case $w \nvDash A$ or $w \vDash B$;
- $w \vDash A \leftrightarrow B$ just in case either $w \nvDash A$ or $w \vDash B$ and either $w \nvDash B$ or $w \vDash A$;
- $w \vDash \Box A$ just in case for every $v \in W$ such that wRv, $v \vDash A$.

Given a model $\mathcal{M}$, we write $\mathcal{M} \vDash A$ to mean that A is true in $\mathcal{M}$. This obtains just in case for every $w \in @, w \vDash A$. In other words, a sentence is true in a model if it is true at each of the actual worlds. We write $\mathcal{M} \vDash \Gamma$ to mean that every member of Γ is true in $\mathcal{M}$.

We must impose constraints on models so that our theory of identity is true in every model. We impose the following two constraints:

- For every $w \in W$, $I(w)$ is an equivalence relation;
- $\langle d, e \rangle \in I(w)$ for every $w \in @$ just in case d is identical to e.

The first constraint is needed for Re, Sy, and Tr. It is a very natural constraint. The second constraint is needed for PII^e. In effect, it says that if two objects are identical at every actual world then they are identical simpliciter (and thus are identical at all worlds). Notice that there is no requirement that distinct worlds otherwise agree about identity.

If we also have PII^{-1e} in the logic, we need to impose a third constraint:

- If $w \vDash E(a)$ and $w \nvDash E(b)$ for every $w \in @$, then $\langle \Phi(a), \Phi(b) \rangle \notin I(w)$ for every $w \in @$.

In other words, if $E(a)$ is true at every actual world and $E(b)$ is false at every actual world then $\neg a = b$ is true at every actual world.

No additional constraints are needed for K and Nec. Additional constraints are needed for the other modal principles. As is familiar, the T, 4, 5, B, and D schemas correspond to the accessibility relation being reflexive, transitive, Euclidean, symmetric, and serial, respectively. If the modal logic is S5, for example, then the accessibility relation is an equivalence relation on worlds.

Notice that the appropriate constraint for S5 is not $R = W \times W$. If a relational model satisfies both the constraint that $R = W \times W$ and the constraint for PII^{-1e} then it obeys PII^c.[33] This is not true for the weaker constraint that R is an equivalence relation.

If we add the rule $A/\Box A$ to the logic, the corresponding constraint on relational models is that $@ = W$. In other words, a sentence is true in a model just in case it is true at all worlds in the model.

---

[33] According to the correspondence between relational and algebraic models defined below, the constraint that $R = W \times W$ corresponds to the constraint that $N(p) = 1$ if $p = 1$ and $N(p) = 0$ otherwise.

Notice that if @ is a singleton in a model, then PII^c will be true in the model: If $x = y$ is false at the actual world, then $x = y \to (F(x) \leftrightarrow F(y))$ is true at the actual world. If $x = y$ is true at the actual world, then x is identical to y and so $x = y \to (F(x) \leftrightarrow F(y))$ is true at the actual world. This shows that it is crucial for our purposes that models can have multiple actual worlds.

As before, we write $\vDash A$ to mean that A is valid. This obtains just in case for every model $\mathcal{M}$, $\mathcal{M} \vDash A$. We write $\Gamma \vDash A$ to mean that Γ entails A. This obtains just in case for every model $\mathcal{M}$, if $\mathcal{M} \vDash \Gamma$ then $\mathcal{M} \vDash A$.

If the language lacks a modal operator, the definition of a model can be simplified in the obvious ways.

5.2. Soundness and Completeness. It is straightforward to prove the strong soundness of each of the weakly classical logics with identity described above with respect to the relational semantics with the relevant constraints imposed on models.

The proof of strong completeness is a bit more complicated. Let $\vdash$ be the consequence relation for any of the weakly classical normal modal logics with identity described above. Let Γ be a set of sentences that is consistent with respect to this consequence relation. We show how to define a model for Γ that can be used in the proof of strong completeness.

In the construction, we make use of a second consequence relation. Let $\vdash^-$ be the consequence relation that has the same axiomatization as $\vdash$ except that it lacks the rules PII^e and PII^{-1e}. For example, if $\vdash$ is the minimal weakly classical consequence relation that obeys Re, Sy, Tr, PII^e, K, Nec, and T, then $\vdash^-$ is the minimal weakly classical consequence relation that obeys Re, Sy, Tr, K, Nec, and T. $\vdash^-$ is always a strongly classical consequence relation.

We say that a set of sentences, Θ, is maximal consistent$^-$ just in case it is maximal consistent with respect to $\vdash^-$. That is, it is consistent with respect to $\vdash^-$ and it has no proper superset that is consistent with respect to $\vdash^-$. Since $\vdash^-$ is a strongly classical consequence relation, for any maximal consistent$^-$ set, Θ, $A \in \Theta$ just in case $\neg A \notin \Theta$; $A \wedge B \in \Theta$ just in case both $A \in \Theta$ and $B \in \Theta$; $A \vee B \in \Theta$ just in case either $A \in \Theta$ or $B \in \Theta$; $A \to B \in \Theta$ just in case either $A \notin \Theta$ or $B \in \Theta$; and $A \leftrightarrow B \in \Theta$ just in case either $A \notin \Theta$ or $B \in \Theta$ and either $B \notin \Theta$ or $A \in \Theta$. Thus, Θ behaves rather like a world.

For any set of sentences in the language, Δ, let Δ^+ be the closure of Δ under $\vdash$. That is, $\Delta^+ = \{A \mid \Delta \vdash A\}$.

Given any set Γ and any necessitation-closed set Δ such that $\Gamma^= \subseteq \Delta \subseteq \Gamma$, we make the following definitions:

$\mathbf{W}^\Delta = \{\Theta \mid \Theta$ is a maximal consistent$^-$ set of sentences containing $\Delta^+\}$;

$\mathbf{R}^\Delta = \{\langle w, v \rangle \mid$ for every sentence A, if $\Box A \in w$ then $A \in v\}$;

$\mathbf{@}^\Gamma = \{\Theta \mid \Theta$ is a maximal consistent$^-$ set of sentences containing $\Gamma^+\}$;

$\mathbf{D}^\Gamma = \{a^\Gamma \mid a$ is a constant in the language$\}$;

$\mathbf{I}^{\Gamma,\Delta}(w) = \{\langle a^\Gamma, b^\Gamma \rangle \mid a = b \in w\}$;

$\mathbf{\Phi}^{\Gamma,\Delta}(a) = a^\Gamma$;

$\mathbf{\Phi}^{\Gamma,\Delta}(P_m^n)(w) = \{\langle a_1^\Gamma, \ldots, a_n^\Gamma \rangle \mid P_m^n a_1 \ldots a_n \in w\}$;

$\mathbf{M}^{\Gamma,\Delta} = \langle \mathbf{W}^\Delta, \mathbf{R}^\Delta, \mathbf{@}^\Gamma, \mathbf{D}^\Gamma, \mathbf{I}^{\Gamma,\Delta}, \mathbf{\Phi}^{\Gamma,\Delta} \rangle$.

If the language lacks a modal operator, the construction can be simplified in the obvious ways.

We can show that if Γ is consistent and Δ is a necessitation-closed set such that $\Gamma^= \subseteq \Delta \subseteq \Gamma$, then $\mathbf{M}^{\Gamma, \Delta}$ is a model. This model satisfies the relevant constraints for the consequence relation $\vdash$ (with the sole exception of the case of the rule $A/\Box A$). $\mathbf{M}^{\Gamma, \Delta} \vDash A$ just in case $\Gamma \vdash A$. Since $\Gamma^=$ is a necessitation-closed set, we can use $\mathbf{M}^{\Gamma, \Gamma^=}$ as our model for Γ in the proof of strong completeness.

If we add the rule $A/\Box A$ to the logic, Γ will be a necessitation-closed set. So we can use the model $\mathbf{M}^{\Gamma, \Gamma}$ in the proof of strong completeness. In this model, the set of actual worlds is the set of worlds, since $@^{\Gamma} = \mathbf{W}^{\Gamma}$.

5.3. The Relationship between the Approaches. There is a natural mapping from relational models to algebraic models. Given any relational model $\mathcal{M} = \langle W, R, @, D, I, \Phi \rangle$, we can define a corresponding algebraic model $\mathcal{M}' = \langle S, \leq, F, N, D, I', \Phi' \rangle$ as follows:

$S = \mathcal{P}(W)$ (that is, the powerset of W);

$p \leq q$ just in case $p \subseteq q$;

$F = \{p \mid @ \subseteq p\}$;

$N(p) = \{w \mid \text{ if } wRv \text{ then } v \in p\}$;

D is unchanged;

$I'(d, e) = \{w \mid \langle d, e \rangle \in I(w)\}$;

$\Phi'(a) = \Phi(a)$;

$\Phi'(P_m^n)(d_1, \ldots, d_n) = \{w \mid \langle d_1, \ldots, d_n \rangle \in \Phi(w)(P_m^n)\}$.

If $\mathcal{M}$ has α worlds for some cardinal α, $\mathcal{M}'$ will have 2^α semantic values. Notice that if $@ = W$, then F will contain only the top value of the Boolean algebra. If $@$ contains only a single world, then F will be an ultrafilter (and since the filter is complete, a principal ultrafilter).

It is straightforward to show that $\mathcal{M} \vDash A$ just in case $\mathcal{M}' \vDash A$. It is also straightforward to show that if $\mathcal{M}$ satisfies the constraints for relational models for any of the weakly classical theories of identity described above, $\mathcal{M}'$ will satisfy the corresponding constraints for algebraic models.

$\mathcal{M}'$ is based on a complete atomic Boolean algebra and has a complete filter. (The atoms of the algebra are the singletons. Given a set of of elements of the algebra, their suprema is their union and their infima is their intersection.) It follows that, in the statement of the strong completeness result for the algebraic semantics, we can require that the Boolean algebra of each model be atomic and complete and that the filter be complete.

There is also a mapping in the other direction. Suppose that we have an algebraic model based on an atomic complete Boolean algebra with a complete filter. By the Stone Representation theorem, the model will be isomorphic to a model based on a powerset algebra. Given such a model $\mathcal{M} = \langle \mathcal{P}(W), \subseteq, F, N, D, I, \Phi \rangle$, we can define a relational model $\mathcal{M}' = \langle W, R, @, D, I', \Phi' \rangle$ as follows:

W is the carrier set of $\mathcal{P}(W)$;

wRv just in case $w \notin N(W - \{v\})$;

@ is the intersection of F;

D is unchanged;

$I(w) = \{\langle d, e \rangle \mid w \in I(d, e)\}$;

$\Phi'(a) = \Phi(a)$;

$\Phi'(P_m^n)(w) = \{\langle d_1, \dots, d_n \rangle \mid w \in \Phi(P_m^n)(d_1, \dots, d_n)\}$.

If F contains only the top value of the Boolean algebra, then @ $= W$. If F is an ultrafilter (and thus a principal ultrafilter), then @ will contain only a single world.

It is straightforward to show that $\mathcal{M} \vDash A$ just in case $\mathcal{M}' \vDash A$. It is also straightforward to show that if $\mathcal{M}$ satisfies the constraints for algebraic models for any of the weakly classical theories of identity described above, $\mathcal{M}'$ will satisfy the corresponding constraints for relational models.

These mappings show that the algebraic and relational approaches to semantics are essentially equivalent, at least if we require that algebraic models be based on complete atomic Boolean algebras and have complete filters. In the remainder of the paper, we'll rely on the algebraic approach to semantics. All of the models used in the proofs below will be finite, so they will satisfy these requirements. Corresponding relational models will be easy to read off from these algebraic models.

§6. The Quasi-Formal Arguments Revisited. We are finally in a position to return to the three quasi-formal arguments for the strictness of identity. Let us discuss them in turn.

6.1. The Collapse Argument. Let us first consider the collapse argument. Recall that the core of the collapse argument is the following technical result: Suppose we have a language with two identity predicates, $=_1$ and $=_2$, that each obey the strongly classical theory of identity. That is, they each obey Re and PII^c. Then we can show that $\vdash a =_1 b \leftrightarrow a =_2 b$. Moreover, we can show that the two identity predicates are fully intersubstitutable.

The situation is different if we move to a weakly classical logic with identity. Consider a language in which we have two identity predicates, $=_1$ and $=_2$, as well as individual constants, predicates, the usual logical connectives, and parentheses. (For simplicity, we operate in a language without a modal operator.) Suppose that the consequence relation is the minimal weakly classical consequence relation such that each identity predicate obeys Re, Sy, Tr, and PII^e. This consequence relation can be axiomatized using any Hilbert-style axiomatization of classical logic with the addition of the following axioms and rules:

$a =_1 a$	Reflexivity$_1$ (Re$_1$)
$a =_2 a$	Reflexivity$_2$ (Re$_2$)
$a =_1 b \to b =_1 a$	Symmetry$_1$ (Sy$_1$)
$a =_2 b \to b =_2 a$	Symmetry$_2$ (Sy$_2$)
$(a =_1 b \wedge b =_1 c) \to a =_1 c$	Transitivity$_1$ (Tr$_1$)
$(a =_2 b \wedge b =_2 c) \to a =_2 c$	Transitivity$_2$ (Tr$_2$)
$a =_1 b, E(a)/E(b)$	Indiscernibility of Identicals$_1^e$ (PII$_1^e$)
$a =_2 b, E(a)/E(b)$	Indiscernibility of Identicals$_2^e$ (PII$_2^e$)

We can assume that the only other rule of inference is Modus Ponens.

The semantics for this language is a simple modification of the algebraic semantics presented above. A model is a septuple, $\langle S, \leq, F, D, I_1, I_2, \Phi \rangle$. We let $\|a =_1 b\| = I_1(\Phi(a), \Phi(b))$ and $\|a =_2 b\| = I_2(\Phi(a), \Phi(b))$. We impose the following constraints on models:

- $I_1(d, e) \in F$ just in case $I_1(d, e) = 1$ just in case d is identical to e;
  $I_2(d, e) \in F$ just in case $I_2(d, e) = 1$ just in case d is identical to e;
- $I_1(d, e) = I_1(e, d)$; $I_2(d, e) = I_2(e, d)$;
- $I_1(d, e) \sqcap I_1(e, f) \leq I_1(d, f)$; $I_2(d, e) \sqcap I_2(e, f) \leq I_2(d, f)$.

It is straightforward to modify the proofs above to show the strong soundness and strong completeness of this consequence relation with respect to the semantics.

PII$_1^c$ and PII$_2^c$ are not theorems of this theory. So the argument for the intersubstitutivity of $=_1$ and $=_2$ does not go through.

We can show that $a =_1 b \dashv\vdash a =_2 b$.[34]

PROOF. By Re$_2$, $\vdash a =_2 a$. By PII$_1^e$, $a =_1 b, a =_2 a \vdash a =_2 b$. So $a =_1 b \vdash a =_2 b$. Similarly, using PII$_2^e$ and Re$_1$, $a =_2 b \vdash a =_1 b$. Therefore, $a =_1 b \dashv\vdash a =_2 b$.[35] $\square$

However, $\neg a =_1 b \nvdash \neg a =_2 b$ and $\neg a =_2 b \nvdash \neg a =_1 b$.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I_1, I_2 \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m\}$; $I_1(l, m) = I_1(m, l) = 0$ and $I_2(l, m) = I_2(m, l) = s$; $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

In this model, $\|\neg a_1 =_1 a_2\| = 1$ but $\|\neg a_1 =_2 a_2\| = -s = t \notin F$. Therefore, $\neg a_1 =_1 a_2 \nvDash \neg a_1 =_2 a_2$. (To show that $\neg a =_2 b \nvDash \neg a =_1 b$, simply switch I_1 and I_2.) $\square$

Thus, the two identity predicates are not fully intersubstitutable.

It follows that $\nvdash a =_1 b \leftrightarrow a =_2 b$. Indeed, the sentence $\neg(a =_1 b \leftrightarrow a =_2 b)$ is consistent.

PROOF. Let $\mathcal{M} = \langle S, \leq, F, D, I_1, I_2 \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m\}$; $I_1(l, m) = I_1(m, l) = s$ and $I_2(l, m) = I_2(m, l) = t$; $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

In this model, $\|\neg(a_1 =_1 a_2 \leftrightarrow a_1 =_2 a_2)\| = -(t \sqcap s) = -0 = 1 \in F$. $\square$

Moreover, the set $\{a =_1 b \vee a =_1 c, \neg(a =_2 b \vee a =_2 c)\}$ is consistent.

---

[34]Thus, on our weakly classical theory of identity, identity is unique in the sense of Belnap (1962). However, as we will see below, this does not entail that $=_1$ and $=_2$ are intersubstitutable in embedded contexts.

[35]There is also a simple semantic proof: $\mathcal{M} \vDash a =_1 b$ just in case $I_1(\Phi(a), \Phi(b)) \in F$ just in case $\Phi(a)$ is identical to $\Phi(b)$ just in case $I_2(\Phi(a), \Phi(b)) \in F$ just in case $\mathcal{M} \vDash a =_2 b$.

PROOF. Let $\mathcal{M} = \langle S, \leq, F, D, I_1, I_2, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m, n\}$; $I_1(l, m) = I_1(m, l) = s$, $I_1(l, n) = I_1(n, l) = t$, $I_1(m, n) = I_1(n, m) = 0$, and $I_2(l, m) = I_2(m, l) = I_2(l, n) = I_2(n, l) = I_2(m, n) = I_2(n, m) = 0$; $\Phi(a_1) = l$, $\Phi(a_2) = m$, and $\Phi(a_3) = n$.

In this model, $\|a_1 =_1 a_2 \vee a_1 =_1 a_3\| = I_1(l, m) \sqcup I_1(l, n) = s \sqcup t = 1 \in F$. $\|\neg(a_1 =_2 a_2 \vee a_1 =_2 a_3)\| = -(I_2(l, m) \sqcup I_2(l, n)) = -(0 \sqcup 0) = -0 = 1 \in F$. $\square$

Thus, it is consistent to accept that the two identity predicates are not equivalent. It is also consistent to accept both a sentence and the negation of the sentence that differs from the original sentence only in that it has occurrences of $=_2$ where the original sentence has occurrences of $=_1$. In short, the two identity predicates may behave very differently in embeddings.

Now suppose that we add the following two additional rules to the proof theory:

$E(a), \neg E(b) / \neg a =_1 b \qquad \text{PII}_1^{-1e}$

$E(a), \neg E(b) / \neg a =_2 b \qquad \text{PII}_2^{-1e}$

It is easy to show that this new consequence relation is strongly sound and strongly complete with our semantics given an additional restriction on models: If $\|E(a)\| \in F$ and $-\|E(b)\| \in F$, then $-I_1(\Phi(a), \Phi(b)) \in F$ and $-I_2(\Phi(a), \Phi(b)) \in F$.

As before, on this consequence relation, $a =_1 b \dashv\vdash a =_2 b$. We can now also show that $\neg a =_1 b \dashv\vdash \neg a =_2 b$.

PROOF. By Re_1, $\vdash a =_1 a$. By PII_2^{-1e}, $a =_1 a, \neg a =_1 b \vdash \neg a =_2 b$. So $\neg a =_1 b \vdash \neg a =_2 b$. By the analogous reasoning, $\neg a =_2 b \vdash \neg a =_1 b$. $\square$

Moreover, the sentence $\neg(a =_1 b \leftrightarrow a =_2 b)$ is no longer consistent.

PROOF. By Re_1 and Re_2, $\vdash a =_1 a$ and $\vdash a =_2 a$. So $\vdash a =_1 a \leftrightarrow a =_2 a$. By PII_1^{-1e}, $a =_1 a \leftrightarrow a =_2 a, \neg(a =_1 b \leftrightarrow a =_2 b) \vdash \neg a =_1 b$. So $\neg(a =_1 b \leftrightarrow a =_2 b) \vdash \neg a =_1 b$. Similarly, using PII_2^{-1e}, $\neg(a =_1 b \leftrightarrow a =_2 b) \vdash \neg a =_2 b$. $\neg a =_1 b, \neg a =_2 b \vdash a =_1 b \leftrightarrow a =_2 b$. Hence, $\neg(a =_1 b \leftrightarrow a =_2 b) \vdash a =_1 b \leftrightarrow a =_2 b$. Therefore, $\neg(a =_1 b \leftrightarrow a =_2 b)$ is inconsistent. $\square$

So we can no longer consistently maintain that the two identity predicates are inequivalent.

Nevertheless, it remains true that $\nvdash a =_1 b \leftrightarrow a =_2 b$.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I_1, I_2, \Phi \rangle$, where $S = \{1, r, s, t, u, v, w, 0\}$ and $\leq$ is the relation depicted in the following Hasse diagram:
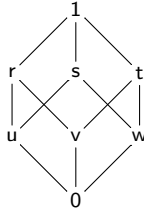
$F = \{1\}$; $D = \{l, m\}$; $I_1(l, m) = I_1(m, l) = r$ and $I_2(l, m) = I_2(m, l) = s$; and $\Phi(a_1) = l$, $\Phi(a_2) = m$, $\Phi(P_1^1)(l) = r$, and $\Phi(P_1^1)(m) = s$.

This model satisfies the additional constraint: Suppose $\|E(a_1)\| = 1$. $E(a_1)$ is equivalent to a sentence in Conjunctive Normal Form – that is, a conjunction of

disjunctions of atomic sentences and negations of atomic sentences. Each of the conjuncts has value 1. By a tedious enumeration of cases, it can be shown that substituting occurrences of a_2 for some of the occurrences of a_1 in a conjunct yields a sentence with value 1, r, or s. Thus, $E(a_2)$ has value 1, r, s, or u. So there is no pair of sentences $E(a_1)$ and $E(a_2)$ such that $E(a_1)$ and $\neg E(a_2)$ both have designated values. By a similar argument, there is no pair of sentences $E(a_1)$ and $E(a_2)$ such that $E(a_2)$ and $\neg E(a_1)$ both have designated values. So the model satisfies the constraint.

In this model, $\|a_1 =_1 a_2 \leftrightarrow a_1 =_2 a_2\| = (-r \sqcap s) \sqcap (-s \sqcup r) = s \sqcap r = u \notin F.$ $\square$

So we can refrain from endorsing the claim that the two identity relations are equivalent. Similarly, $a =_1 b \vee a =_1 c$ may be true in a model without $a =_2 b \vee a =_2 c$ being true in the model. So it is consistent to accept $a =_1 b \vee a =_1 c$ without also accepting $a =_2 b \vee a =_2 c$. Thus, even if we add the rules PII_1^{-1e} and PII_2^{-1e} to the logic, the two identity predicates are not fully intersubstitutable.

Given these results, what should we make of the collapse argument and its alleged consequences? The weakly classical theory of identity generated by Re, Sy, Tr and PII^e is a very natural theory. It contains the claim that the identity relation is an equivalence relation and the rule PII^e. If we adopt this theory as our theory of identity (perhaps also including the rule PII^{-1e}), we can no longer make use of the collapse argument in arguing that: (i) the identity predicates of distinct speakers must have the same extension; (ii) the identity predicate determinately refers to a single determinate relation; and (iii) identities are necessary.

We can still argue that the identity predicates of two speakers must agree at least insofar as $a =_1 b \dashv\vdash a =_2 b$. In other words, if one speaker accepts an identity statement and the other does not, at least one of them is making a mistake. However, if identity theory does not contain PII^{-1e}, the speakers need not agree about which identity statements they deny. Moreover, even if identity theory includes PII^{-1e}, the speakers needn't agree about how the identity predicate behaves in other embedded contexts. For instance, one speaker may accept that $a = b \vee a = c$ while the other may not. Thus, we can no longer use the collapse argument to argue that two speakers must have identity predicates with the same extension.

Similarly, we can still argue that any two precisifications of the identity relation must agree at least insofar as $a =_1 b \dashv\vdash a =_2 b$. In other words, if an identity is true on one precisification, it must be true on the other. But if identity theory does not include PII^{-1e}, precisifications need not agree about when the negation of an identity claim is true. Moreover, even if identity theory includes PII^{-1e}, precisifications need not agree on how the identity predicate behaves in other embedded contexts. For instance, the sentence $a = b \vee a = c$ may be true on one precisification but not on another. Thus, two precisifications of the identity predicate may behave differently. We can no longer use the collapse argument to argue that identity cannot be vague or otherwise indeterminate.

Finally, we can still argue that if $a = b$ obeys our weakly classical theory of identity, so does $\square a = b$. We can still argue that $a = b$ and $\square a = b$ must agree at least insofar as $a = b \dashv\vdash \square a = b$. But irrespective of whether we include PII^{-1e} in our theory of identity, $a = b$ and $\square a = b$ may behave very differently

in embeddings. In particular, we need not accept that $\vdash a = b \to \Box a = b$. So we can no longer make use of the collapse argument to argue that if an identity obtains, it obtains necessarily.

6.2. Modality. Let us now consider the modal argument. Recall that the core of the modal argument is the following technical result: Suppose we have a strongly classical normal modal logic with identity. Suppose our theory of identity contains Re and PII^c. Then we can show that $\vdash a = b \to \Box a = b$. Moreover, given additional modal principles, we can show that $\vdash a = b \leftrightarrow \Box a = b$, $\vdash \neg a = b \leftrightarrow \Box \neg a = b$, and $\vdash \Box a = b \lor \Box \neg a = b$. In particular, these results obtain in S5, the best candidate for the logic of metaphysical necessity. This suggests that identity and distinctness obtain necessarily and that given any pair of objects, they are either necessarily identical or necessarily distinct. There is no room for any sort of contingency concerning identity or distinctness.

The situation is different in weakly classical logic. Since PII^c does not obtain, the argument that $\vdash a = b \to \Box a = b$ does not go through. It is true that in our base weakly classical normal modal logic with identity, $a = b \vdash \Box a = b$.

PROOF. By Re, $\vdash a = a$. By PII^e, $a = b, \Box a = a \vdash a = a$. By Nec, $\vdash \Box a = a$. Therefore, $a = b \vdash \Box a = b$. $\qquad\square$

In addition, if we have the T principle in the logic, we can show that $\vdash \Box a = b \to a = b$. Thus, in KT, we can show that $a = b \dashv\vdash \Box a = b$. Using T, we can also show that $\vdash \Box \neg a = b \to \neg a = b$ and thus that $\Box \neg a = b \vdash \neg a = b$. However, even if we accept the S5 principles for necessity, $\neg a = b \nvdash \Box \neg a = b$, $\neg \Box \neg a = b \nvdash a = b$, and $\nvdash \Box a = b \lor \Box \neg a = b$. Indeed, it is consistent to accept $\neg(\Box a = b \lor \Box \neg a = b)$.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, N, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1, s\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = t$; $N(1) = 1$ and $N(0) = N(s) = N(t) = 0$; and $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

In this model, $\|\neg a_1 = a_2\| = -I(\Phi(a_1), \Phi(a_2)) = -I(l, m) = -t = s \in F$ but $\|\Box \neg a_1 = a_2\| = N(s) = 0 \notin F$.

In this model, $\|\neg \Box \neg a_1 = a_2\| = -N(-I(l, m)) = -N(-t) = -N(s) = -0 = 1 \in F$ but $\|a_1 = a_2\| = I(l, m) = t \notin F$.

In this model, $\|\Box a_1 = a_2 \lor \Box \neg a_1 = a_1\| = N(I(l, m)) \sqcup N(-I(l, m)) = N(t) \sqcup N(-t) = N(t) \sqcup N(s) = 0 \sqcup 0 = 0 \notin F$. $\|\neg(\Box a_1 = a_2 \lor \Box \neg a_1 = a_1)\| = 1 \in F$. $\qquad\square$

Moreover, $\nvdash a = b \to \Box a = b$.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, N, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1, s\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = s$; $N(1) = 1$ and $N(0) = N(s) = N(t) = 0$; and $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

In this model, $\|a_1 = a_2 \to \Box a_1 = a_2\| = -I(\Phi(a_1), \Phi(a_2)) \sqcup N(I(\Phi(a_1), \Phi(a_2))) = -s \sqcup N(s) = t \sqcup 0 = t \notin F$. $\qquad\square$

If we accept a weakly classical normal modal logic with identity, we cannot both accept an identity and the claim that the identity is not necessary. That is, the sentence $a = b \land \neg \Box a = b$ is inconsistent, as is the logically equivalent sentence $\neg(a = b \to \Box a = b)$. However, we can accept that it is possible that an

identity is contingently true. That is, $\neg\Box\neg(a = b \wedge \neg\Box a = b)$ is consistent, as is the logically equivalent sentence $\neg\Box(a = b \to \Box a = b)$. So we can consistently accept a positive claim about how $a = b \to \Box a = b$ is defective, and not just withhold acceptance of it.

PROOF. Let $\mathcal{M} = \langle S, \leq, F, D, I, N, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1, s\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = t$; $N(1) = 1$ and $N(0) = N(s) = N(t) = 0$; and $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

In this model, $\|\neg\Box\neg(a_1 = a_2 \wedge \neg\Box a_1 = a_2)\| = -N(-(I(\Phi(a_1), \Phi(a_2)) \sqcap -N(I(\Phi(a_1), \Phi(a_2)))))) = -N(-(I(l, m) \sqcap -N(I(l, m)))) = -N(-(t \sqcap -N(t))) = -N(-(t \sqcap -0)) = -N(-(t \sqcap 1)) = -N(-t) = -N(s) = -0 = 1 \in F$. $\qquad\Box$

All of these results continue to hold even if we add the rule PII^{-1e} to our logic. Of the three countermodels presented above, the first and the third satisfy the constraint that if $\|E(a)\| \in F$ and $-\|E(b)\| \in F$ then $-I(\Phi(a), \Phi(b)) \in F$. (In each of these models, the only members of the domain are l and m and $-I(l, m) \in F$.) The result that $\nvdash a = b \to \Box a = b$ requires a more complex model.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, N, \Phi \rangle$, where $S = \{1, r, s, t, u, v, w, 0\}$ and $\leq$ is the relation depicted in the following Hasse diagram:
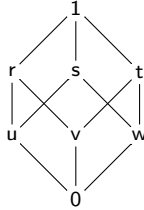
$F = \{1\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = r$; $N(1) = 1, N(r) = N(s) = N(u) = u$, and $N(0) = N(t) = N(v) = N(w) = 0$; and $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

This model satisfies the additional constraint: Suppose $\|E(a_1)\| = 1$. Since we are operating in S5, $E(a_1)$ is equivalent to a sentence in Modal Conjunctive Normal Form – that is, a conjunction of disjunctions of sentences of the form A, $\Box A$, and $\neg\Box A$, where A has no modal operators.[36] Each of the conjuncts has value 1. By a tedious enumeration of cases, it can be shown that substituting occurrences of a_2 for some of the occurrences of a_1 in a conjunct yields a sentence with value 1, r, s, or u. Thus, $E(a_2)$ has value 1, r, s, or u. So there is no pair of sentences $E(a_1)$ and $E(a_2)$ such that $E(a_1)$ and $\neg E(a_2)$ both have designated values. By a similar argument, there is no pair of sentences $E(a_1)$ and $E(a_2)$ such that $E(a_2)$ and $\neg E(a_1)$ both have designated values. So the model satisfies the constraint.

In this model, $\|a_1 = a_2 \to \Box a_1 = a_2\| = -r \sqcap N(r) = w \sqcap u = s \notin F$. $\qquad\Box$

We do get an additional result by adding the rule PII^{-1e}. Namely, $\neg\Box a = b \vdash \neg a = b$.[37]

---

[36]See Hughes & Cresswell (1996), page 82.
[37]I owe this observation to Heck (1998), page 288.

PROOF. Here is a countermodel for the case that the logic does not contain PII^{-1e}: Let $\mathcal{M} = \langle S, \leq, F, D, I, N, \Phi \rangle$, where $S = \{1, s, t, 0\}$; $\langle S, \leq \rangle$ is the four element Boolean algebra such that $0 < s, t < 1$; $F = \{1\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = s$; $N(1) = 1$ and $N(0) = N(s) = N(t) = 0$; and $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

In this model, $\|\neg\Box a_1 = a_2\| = -N(I(\Phi(a_1), \Phi(a_2))) = -N(s) = -0 = 1 \in F$. $\|\neg a_1 = a_2\| = -I(\Phi(a_1), \Phi(a_2)) = -s = t \notin F$.

Now suppose that the logic contains PII^{-1e}. By Re and Nec, $\vdash \Box a = a$. By PII^{-1e}, $\Box a = a, \neg\Box a = b \vdash \neg a = b$. So $\neg\Box a = b \vdash \neg a = b$. $\square$

The claim that $\neg\Box a = b \vdash \neg a = b$ is the contrapositive of $a = b \vdash \Box a = b$. It should not be surprising, then, that we get the contrapositive of $a = b \vdash \Box a = b$ when we add the contrapositive of PII^e to the logic.

Given these results, what should we conclude about the necessity of the identity and distinctness relations? The weakly classical theory containing Re, Sy, Tr, and PII^e is very natural. If we adopt it as our theory of identity, and we accept the S5 principles for $\Box$, we can no longer argue that the identity and distinctness relations obtain necessarily. It is true that we cannot accept an identity and fail to accept that it is necessary. But we can accept a distinctness claim without accepting that it is necessary. Moreover, we need not accept the conditional $a = b \rightarrow \Box a = b$. Indeed, we can accept $\neg\Box\neg(a = b \wedge \neg\Box a = b)$. That is, we can accept that it is possible for an identity to be true without being necessarily true. More strikingly still, it is consistent to accept that neither an identity nor its negation is necessary. And these results hold even if we endorse the rule PII^{-1e}.[38] This suggests there is room for contingency in both the identity and the distinctness relations.

6.3. Vagueness. Finally, let us consider the vagueness argument. The operator Det can be treated as a modal operator, with the same semantics as $\Box$. As is true for the case of $\Box$, if Det obeys a weakly classical normal modal logic with identity, we can show that $a = b \vdash \text{Det}\, a = b$. If, in addition, $\text{Det}\, A \vdash A$, we can show $a = b \dashv\vdash \text{Det}\, a = b$. However, even if our modal logic for Det is S5 (or some weakening of it), $\neg a = b \nvdash \text{Det}\,\neg a = b$, $\neg\text{Det}\,\neg a = b \nvdash a = b$, $\nvdash \text{Det}\, a = b \vee \text{Det}\,\neg a = b$, and $\nvdash a = b \rightarrow \text{Det}\, a = b$. Indeed, it is consistent to accept $\neg(\text{Det}\, a = b \vee \text{Det}\,\neg a = b)$. Moreover, while it is not consistent to accept $a = b \wedge \neg\text{Det}\, a = b$, it is consistent to accept $\neg\text{Det}\,\neg(a = b \wedge \neg\text{Det}\, a = b)$ and the logically equivalent $\neg\text{Det}(a = b \rightarrow \text{Det}\, a = b)$. And each of these results continues to obtain even if we add the rule PII^{-1e} to the theory of identity.

Suppose $\text{V}\, A$ is an abbreviation for $\neg\text{Det}\, A \wedge \neg\text{Det}\,\neg A$. (Thus, $\text{V}\, A$ is logically equivalent to $\text{V}\,\neg A$.) Since $a = b \vdash \text{Det}\, a = b$, we have it that $a = b \vdash \neg\text{V}\, a = b$. However, we can easily show that $\neg a = b \nvdash \neg\text{V}\, a = b$. Moreover, $\nvdash a = b \rightarrow \neg\text{V}\, a = b$, $\nvdash \neg a = b \rightarrow \neg\text{V}\, a = b$, $\nvdash \text{V}\, a = b \rightarrow \neg a = b$, and $\text{V}\, a = b \nvdash \neg a = b$. Indeed, $\text{V}\, a = b$ and the equivalent $\text{V}\,\neg a = b$ are each consistent.

Given this weakly classical theory of identity, Evans's argument for the determinacy of identity breaks down. We no longer have PII^c, and so cannot use it to

---

[38]If we endorse PII^{-1e} and we claim of some identity that neither it nor its negation is necessary, we will be committed to the negation of the identity. But we won't be committed to the necessity of the negation.

show that $\vdash \mathrm{V}\, a = b \to \neg a = b$. Indeed, as we have seen, this sentence is no longer a theorem. Moreover, since $\nvdash \neg a = b \to \neg \mathrm{V}\, a = b$ and $\nvdash \neg a = b \to \neg \mathrm{V}\, \neg a = b$, the later parts of Evans's argument fail, too.

The second strategy for arguing for the determinacy of identity and distinctness – based on an application of the Barcan-Kripke argument – also fails. It fails in exactly the same way as does the argument for the necessity of identity and distinctness.

So far, this is all straightforward. There is a complication, however. Many have found the rule $A/\operatorname{Det} A$ to be plausible. Indeed, this is a valid rule on certain supervaluational treatments of vagueness.[39] On our algebraic semantics, this rule corresponds to the constraint that the only designated value in a model is the top value of the Boolean algebra. The trouble is that if our logic contains this rule as well as PII^{-1e}, we can show that $\mathrm{V}\, a = b$ is inconsistent.[40]

PROOF. By our previous results, given PII^{-1e}, $\neg \operatorname{Det} a = b \vdash \neg a = b$. Since $A \vdash \operatorname{Det} A$, we have it that $\neg a = b \vdash \operatorname{Det} \neg a = b$. Thus, $\neg \operatorname{Det} a = b \vdash \operatorname{Det} \neg a = b$. So, $\neg \operatorname{Det} a = b, \neg \operatorname{Det} \neg a = b \vdash \operatorname{Det} \neg a = b \wedge \neg \operatorname{Det} \neg a = b$. It follows that $\mathrm{V}\, a = b \vdash \operatorname{Det} \neg a = b \wedge \neg \operatorname{Det} \neg a = b$. Therefore, $\mathrm{V}\, a = b$ is inconsistent. $\square$

This is the residual force of Evans's original argument.

What should we make of this result? The inconsistency proof requires both the rule PII^{-1e} and the rule $A/\operatorname{Det} A$. So we might consider giving up one of these two rules. As we have seen, PII^{-1e} corresponds to an unattractive constraint in the semantics. Yet, it is a plausible rule governing identity. So it seems difficult to give up this rule. The real issue concerns the rule $A/\operatorname{Det} A$. While this rule is widely endorsed, it does not strike me as obviously correct. Why think that someone committed to A is thereby committed to the determinacy of A?

Nevertheless, let us suppose that both rules are correct. Does this show that we must claim that identity claims are never vague? The answer is no. Since we are operating in a weakly classical logic, Reductio is not a valid meta-rule. So we cannot use it to conclude that $\vdash \neg \mathrm{V}\, a = b$.[41] Indeed, we can show that $\nvdash \neg \mathrm{V}\, a = b$, or equivalently, that $\nvdash \operatorname{Det} a = b \vee \operatorname{Det} \neg a = b$. We can also show that $\nvdash a = b \to \operatorname{Det} a = b$ and $\nvdash \neg a = b \to \operatorname{Det} \neg a = b$.

PROOF. Here is a countermodel: Let $\mathcal{M} = \langle S, \leq, F, D, I, N, \Phi \rangle$, where $S = \{1, r, s, t, u, v, w, 0\}$ and $\leq$ is the relation depicted in the following Hasse diagram:

_____
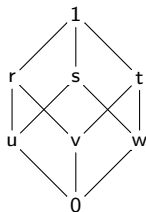
[39]See, for instance, Fine (1975). On a supervaluational treatment, if logical consequence is defined as global validity, then A entails $\operatorname{Det} A$. This is not so if logical consequence is defined as local validity. See Williamson (1994), chapter 5.

[40]See Heck (1998), page 288.

[41]See Heck (1998), page 291.

$F = \{1\}$; $D = \{l, m\}$; $I(l, m) = I(m, l) = r$; $N(1) = 1, N(r) = N(s) = N(u) = u$, and $N(0) = N(t) = N(v) = N(w) = 0$; and $\Phi(a_1) = l$ and $\Phi(a_2) = m$.

As we have already shown (on page 33), this model satisfies the constraint for PII^{-1e}. This model has only a single designated value.

In this model, $\|\text{Det}\, a_1 = a_2 \vee \text{Det}\, \neg a_1 = a_2\| = N(I(l, m)) \sqcup N(-I(l, m)) = N(r) \sqcup N(-r) = N(r) \sqcup N(w) = u \sqcup 0 = u \notin F$.

In this model, $\|a_1 = a_2 \rightarrow \text{Det}\, a_1 = a_2\| = -I(l, m) \sqcup N(I(l, m)) = -r \sqcup N(r) = w \sqcup u = s \notin F$.

In this model, $\|\neg a_1 = a_2 \rightarrow \text{Det}\, \neg a_1 = a_2\| = I(l, m) \sqcup N(-I(l, m)) = r \sqcup N(-r) = r \sqcup N(w) = r \sqcup 0 = r \notin F$. $\square$

So we can withhold assent to the claim that an identity is not vague. We can also withhold assent to the definiteness of identity and the definiteness of non-identity. We just cannot consistently claim that an identity is vague. This provides us with less wiggle room than before, but it still provides us enough to resist endorsing the definiteness or determinacy of the identity and distinctness relations.

Given that we can avoid endorsing the determinacy of identity, is there any positive claim we can consistently accept about how $\neg \text{V}\, a = b$ is defective? For example, can we consistently accept $\neg \text{Det}\, \neg \text{V}\, a = b$? The answer is that we cannot, at least assuming that we are operating in S5. In S5, iterated modalities collapse. However, there are reasons to think that S5 is not the correct modal logic for the definiteness or determinately operators. If higher-order indeterminacy is a genuine phenomenon, then we should move to a weaker modal logic for Det. And if we endorse a weaker modal logic, we may be able to consistently claim that $\neg \text{V}\, a = b$ is defective, and not merely fail to endorse it.

Given these results, what should we conclude about the determinacy of the identity and distinctness relations? The weakly classical theory containing Re, Sy, Tr, and PII^e is a very natural theory. It is plausible that the correct logic for determinacy (or definiteness) is some weakening of S5. If we adopt this theory as our theory of identity, and we adopt some weakening of S5 as our theory of determinacy, we can no longer argue that the identity and distinctness relations are determinate. It is true that we cannot accept an identity and fail to accept that it is determinate. But we can accept a distinctness claim without accepting that it is determinate. Moreover, we need not accept the conditional $a = b \rightarrow \text{Det}\, a = b$. Indeed, we can accept $\neg \text{Det}(a = b \rightarrow \text{Det}\, a = b)$. That is, we can accept that it is not determinate that if an identity is true, it is determinately true. More strikingly still, we can accept that it is vague whether an identity obtains. And these results obtain even if we endorse the rule PII^{-1e}.

If we also endorse the (perhaps dubious) rule $A / \text{Det}\, A$, we can show that $\neg a = b \vdash \text{Det}\, \neg a = b$. Moreover, it is no longer consistent to claim that it is vague whether an identity obtains. Nevertheless, we need not accept $a = b \rightarrow \text{Det}\, a = b$ or $\neg a = b \rightarrow \text{Det}\, \neg a = b$. So the arguments for the determinacy of identity and non-identity fail. Even with this rule, then, there is room to avoid endorsing the determinacy of the identity and distinctness relations.

§7. Discussion. Let's take stock. If we accept strongly classical logic and a theory of identity that contains the claim that identity is reflexive and the claim PII^c, there are compelling arguments that suggest that (i) there is a single identity relation and a single distinctness relation; (ii) the identity and distinctness relations obtain necessarily; and (iii) the identity and distinctness relations obtain definitely and determinately. On the other hand, if we accept a merely weakly classical logic and the weaker theory of identity that contains the claim that identity is an equivalence relation as well as the rule PII^e (and perhaps also the rule PII^{-1e}), the arguments are blocked.

In the weaker theory, we can still argue that $a =_1 b \dashv\vdash a =_2 b$. Given two identity relations, if one accepts an identity claim, so must the other. If we accept PII^{-1e}, we can also argue that $\neg a =_1 b \dashv\vdash \neg a =_2 b$. Given any two identity relations, if one accepts a distinctness claim, so must the other. But even if we accept PII^{-1e}, two identity relations may behave very differently in other embedded contexts. In short, there is room for multiple extensionally-inequivalent identity and distinctness relations.

In the weaker theory of identity, we can argue that $a = b \vdash \Box a = b$ and $a = b \vdash \text{Det}\, a = b$. If we accept an identity claim, we must accept that it is necessary, definite, and determinate. But even if we accept PII^{-1e}, $\neg a = b \nvdash \Box \neg a = b$ and $\neg a = b \nvdash \text{Det}\, \neg a = b$. If we accept a distinctness claim, we need not accept that it is necessary, definite, or determinate. Moreover, $\nvdash a = b \to \Box a = b$ and $\nvdash a = b \to \text{Det}\, a = b$. Indeed, it is consistent to accept $\neg(\Box a = b \vee \Box \neg a = b)$ and $\neg(\text{Det}\, a = b \vee \text{Det}\, \neg a = b)$. That is, we can accept that neither an identity nor its negation is necessary. We can also accept that neither an identity not its negation is definite or determinate. In other words, there is room for contingency, vagueness, and indeterminacy in the identity and distinctness relations.

If we endorse both PII^{-1e} and $A/\text{Det}\, A$, our wiggle room declines further. But even in this setting, we need not accept $a = b \to \text{Det}\, a = b$ or $\neg a = b \to \text{Det}\, \neg a = b$. So we need not accept the definiteness and determinacy of the identity and distinctness relations.

It seems, then, that we can accept a reasonable theory of identity and reasonable principles for $\Box$ and Det without committing ourselves to the uniqueness, necessity, or determinacy of the identity and distinctness relations.

Supposing that our only two options are (a) the strongly classical theory of identity containing Re and PII^c; and (b) the weakly classical theory of identity containing Re, Sy, Tr, and PII^e (and perhaps also PII^{-1e}), how should we decide between them? There are really two decisions to make. The first decision is between strongly classical logic and (merely) weakly classical logic. Assuming we go for weakly classical logic, the second decision is between the two theories of identity. (Given strongly classical logic, the two theories are equivalent.)

How should we decide between strongly classical and weakly classical logic? This question is bound up with fundamental questions concerning the nature of logic. On one picture, logic is intimately tied to reasoning; logic tells us how we ought to (deductively) reason. This picture sits well with the idea that the standard Natural Deduction rules – including the meta-rules – are of central

importance.[42] Perhaps such rules are built into the logical concepts, as certain versions of Conceptual Role Semantics maintain. Even if they are not, such rules tell us how we should reason. Given this picture, we should presumably endorse the classical meta-rules. We should accept strongly classical logic.[43]

On an alternative picture of the nature of logic, logic is not directly connected with reasoning. It is not a normative but a descriptive theory. Logic is the theory of the (logical) consequence relation. It tells us what follows from what.[44] This picture sits well with the idea that what is central to logic is not the Natural Deduction rules but the logical truths and entailment claims. Given this picture, we should presumably endorse (merely) weakly classical logic. We should only go on to accept strongly classical logic if the rules governing the vocabulary in our language permit it.[45]

The choice between strongly classical and weakly classical logic thus depends (in part) on which of these two pictures of the nature of logic is correct.

There is another – semantic – way to think about the choice between strongly classical and weakly classical logic.[46] Consider the claim $\vDash A \to B$. On a simple semantics for strongly classical logic, $\vDash A \to B$ is true just in case whenever A is designated in a model, so is B. In contrast, on our algebraic semantics for weakly classical logic with identity, the claim is true just in case in every model $\|A\| \leq \|B\|$. So we can think about the choice between strongly classical and weakly classical logic by considering the partial order of semantic values. Does this partial order correspond to anything real? If the partial order corresponds to something real, should the semantics for the logical connectives make use of it, or should it only make use of the distinction between designated and undesignated values? These questions do not have obvious answers, but they provide a way to get a handle on the choice between strongly classical and weakly classical logic.

Supposing we end up endorsing weakly classical logic, how should we decide between the two theories of identity? Unfortunately, it is not entirely clear how to make this choice. Both theories are very natural. Perhaps there is some intuitive support for PII^c. It is intuitive that if a is identical to b then a and b share the very same properties. This claim is most naturally understood as a conditional and not as an entailment. But, the difference is subtle, and this is, at best, very weak evidence. I suspect that the best way to decide between

---

[42]Indeed, Gentzen (1935) motivates the natural deduction rules in part by connecting them to reasoning.

[43]This conclusion is not forced, however. We can accept the primacy of Natural Deduction formulations of logic and also accept a merely weakly classical logic. In particular, we can place side conditions on the standard Natural Deduction rules. For instance, we might accept a version of Conditional Introduction restricted so that the rule PII^e cannot be used in sub-derivations. This is broadly analogous to what advocates for supervaluational semantics for truth sometimes do. They accept a version of Conditional Introduction restricted so that the rules governing the truth predicate cannot be used in subderivations. See, for instance, McGee & McLaughlin (2004), section 3.

[44]See Harman (1995) for arguments for this picture of logic.

[45]This conclusion is also not forced. We can accept that logic is the descriptive theory of the consequence relation and still claim that we should endorse strongly classical logic no matter what is in our language. For example, we might claim that the the meta-rules encode important patterns among entailments.

[46]For discussion of this issue, I am grateful to Gabriel Uzquiano.

these two views is to think about the consequences of the decision. Is it plausible that the identity and distinctness relations are necessary? Is it plausible that the identity and distinctness relations are definite and determinate? There may be good reasons from metaphysics to want there to be room for contingency, vagueness, or indeterminacy of identity. Perhaps, for instance, we should accept that there is "vagueness in the world" and perhaps this requires vagueness in the identity relation. If so, assuming that we want to hold onto classical logic, the best alternative may be to endorse a weakly classical theory of identity. At the very least, this should be one of the theoretical options open to us.

At the end of the day, we may decide that the weakly classical theory of identity is insufficiently powerful. We may find reason to endorse the strongly classical theory of identity. Even if we do so, I take it that the results presented here are of interest. They demonstrate that there is a significant difference between strongly classical and (merely) weakly classical logic. They show that there is a significant difference between the Principle of the Indiscernibility of Identicals in conditional form and in rule form. The arguments for the strictness of identity crucially rely on these differences. They rely on the power of strongly classical logic and the power of the Principle of the Indiscernibility of Identicals in conditional form. Even if we ultimately conclude that we should continue to accept the strongly classical theory of identity, it is important that we understand exactly what it is that we are accepting.

REFERENCES

Akiba, K. (2004). Vagueness in the world. *Noûs* **38**(3), 407–429.

Barcan, R. (1947). The identity of individuals in a strict functional calculus of second order. *The Journal of Symbolic Logic* **12**(1), 12–15.

Barnes, E. (2009). Indeterminacy, identity and counterparts: Evans reconsidered. *Synthese* **108**(1), 81–96.

Belnap, N. D. (1962). Tonk, plonk and plink. *Analysis* **22**(6), 130–134.

Evans, G. (1978). Can there be vague objects? *Analysis* **38**(4), 208.

Field, H. (2008). *Saving Truth from Paradox*. Oxford: Oxford University Press.

Fine, K. (1975). Vagueness, truth and logic. *Synthese* **30**(3–4), 265–300.

Gentzen, G. (1934/1935). Untersuchungen ber das logische schließen. *Mathematische Zeitschrift* **39**, 176–210 and 405–431.

Harman, G. (1995). Rationality. In Smith, E. & Osherson, D., editors, *Thinking: An Invitation to Cognitive Science* (Second ed.)., pp. 175–211. Cambridge, MA.: The MIT Press.

Heck, R. (1998). That there might be vague objects (so far as concerns logic). *The Monist* **81**(2), 274–296.

Hughes, G. E., & Cresswell, M. J. (1996). *A New Introduction to Modal Logic*. New York: Routledge.

Kripke, S. (1971). Identity and necessity. In Munitz, M., editor, *Identity and Individuation*, pp. 135–164. New York: New York University Press.

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Wiley-Blackwell.

Lewis, D. (1988). Vague identity: Evans misunderstood. *Analysis* **48**(3), 128–130.

McGee, V., & McLaughlin, B. P. (2004). Logical commitment and semantic indeterminacy: A reply to Williamson. *Linguistics and Philosophy* **27**(1), 123–136.

Putnam, H. (1983). Vagueness and alternative logic. *Erkenntnis* **19**, 297–314.

Quine, W. V. O. (1961). Reply to Professor Marcus. *Synthese* **13**(4), 323–330.

Rasiowa, H., & Sikorski, R. (1970). *The Mathematics of Metamathematics* (Third ed.). Warsaw: Polish Scientific Publishers.

Salmon, N. (1981). *Reference and Essence*. Princeton: Princeton University Press.

Schechter, J. (2011). Juxtaposition: A new way to combine logics. *The Review of Symbolic Logic* **4**(4), 560–606.

Williams, J. R. G. (2008). Multiple actualities and ontically vague identity. *The Philosophical Quarterly* **58**(230), 134–154.

Williamson, T. (1994). *Vagueness*. New York: Routledge.

Williamson, T. (2007). Absolute identity and absolute generality. In Rayo, A. & Uzquiano, G., editors, *Absolute Generality*, pp. 369–390. Oxford: Oxford University Press.

Wright, C. (2001). On being in a quandry: Relativism, vagueness, logical revisionism. *Mind* **110**, 45–98.