This is an electronic version of a paper published in *Review of Symbolic Logic*,
http://dx.doi.org/10.1017/S1755020313000312
Published online: 14 October 2013

# NAIVE TRUTH AND RESTRICTED QUANTIFICATION: SAVING TRUTH A WHOLE LOT BETTER

HARTRY FIELD
New York University and University of Birmingham

**Abstract.** Restricted quantification poses a serious and under-appreciated challenge for nonclassical approaches to both vagueness and the semantic paradoxes. It is tempting to explain "All $A$ are $B$" as "For all $x$, if $x$ is $A$ then $x$ is $B$"; but in the nonclassical logics typically used in dealing with vagueness and the semantic paradoxes (even those where 'if ... then' is a special conditional not definable in terms of negation and disjunction or conjunction), this definition of restricted quantification fails to deliver important principles of restricted quantification that we'd expect. If we're going to use a nonclassical logic, we need one that handles restricted quantification better.

The challenge is especially acute for naive theories of truth—roughly, theories that take $True(\langle A\rangle)$ to be intersubstitutable with $A$, even when $A$ is a "paradoxical" sentence such as a Liar-sentence. A naive truth theory inevitably involves a somewhat nonclassical logic; the challenge is to get a logic that's compatible with naive truth *and also validates intuitively obvious claims involving restricted quantification* (for instance, "If $S$ is a truth stated by Jones, and every truth stated by Jones was also stated by Smith, then $S$ is a truth stated by Smith"). No extant naive truth theory even comes close to meeting this challenge, including the theory I put forth in *Saving Truth from Paradox*. After reviewing the motivations for naive truth, and elaborating on some of the problems posed by restricted quantification, I will show how to do better. (I take the resulting logic to be appropriate for vagueness too, though that goes beyond the present paper.)

In showing that the resulting logic is adequate to naive truth, I will employ a somewhat novel fixed point construction that may prove useful in other contexts.

## Part 1. INTRODUCTION

**§1. Restricted quantification.** Here are some laws we would expect universal restricted quantification ("All $A$ are $B$") to obey:

**(I):** (For all $x$:) If all $A$ are $B$, and $x$ is $A$, then $x$ is $B$.
[Hence presumably: **(I$^-$)** If all $A$ are $B$, and everything is $A$, then everything is $B$.]

**(II):** If everything is $B$, then all $A$ are $B$

**(III):** If all $A$ are $B$, and all $B$ are $C$, then all $A$ are $C$

**(III$_{var}$):** If all $A$ are $B$, and not all $A$ are $C$, then not all $B$ are $C$

**(IV):** If all $A$ are $B$, and all $A$ are $C$, then all $A$ are both $B$ and $C$

**(IV$_{var}$):** If all $A$ are $B$, and not all $A$ are both $B$ and $C$, then not all $A$ are $C$

**(V):** If not all $A$ are $B$, then something is both $A$ and not $B$

**(VI):** If something is both $A$ and not $B$, then not all $A$ are $B$.

Received: May 21, 2013.



doi:10.1017/S1755020313000312

I'm inclined to add that restricted quantification contraposes:

**(C):** All $A$ are $B$ if and only if all not-$B$ are not-$A$.

But even if one is suspicious of this in general, the principles that can be derived by applying **(C)** to the above principles seem highly compelling. For instance, by taking instances of **(I)** and **(II)**, applying **(C)**, and using the redundancy of double negation, one gets the following:

**($\mathbf{I}_c$):** If all $C$ are $D$, and $x$ is not $D$, then $x$ is not $C$

**($\mathbf{II}_c$):** If nothing is $C$ then all $C$ are $D$.

I think these ought to be on our "wish list" even if **(C)** isn't assumed.

All of these laws are of course validated in classical logic, when restricted quantification is handled in the standard way. But in a "naive theory of truth" (to be defined in the next section) the logic can't be fully classical, and the challenge is to get a nonclassical logic that is both adequate to naive truth and that validates these laws. I'll also assume that 'if ... then' obeys modus ponens, and other obvious laws such as that from $A$ and $\neg B$ one can infer that it is not the case that if $A$ then $B$. I also require that the logic allow reasoning by cases, taken to include $\exists$-Elimination as well as $\vee$-Elimination. And I'll assume that the logic includes substitutivity of identity and the standard structural rules, including the transitivity of consequence. (I don't mean to suggest that these assumptions are beyond discussion, but here is not the place.)

I assumed until recently that in order to maintain "naive truth" (to be explained shortly) I was going to have to qualify some of these laws slightly, so that the challenge would simply be to get a logic for naive truth that gave *a pretty good approximation to* these laws. That would be far more than one gets in extant theories of truth (including the theory in Field, 2008). But to my surprise I discovered that it's possible to get all of the laws above (including **(C)**) to hold unrestrictedly in a naive theory of truth.[1]

For handling vagueness as well as for handling truth, there is reason to take nonclassical logic seriously, and there's also reason to want the same laws of restricted quantification. It turns out that this is problematic there too: the nonclassical logics most often discussed for vagueness are (i) Kleene logics, which have no remotely reasonable notion of restricted quantification; and (ii) Lukasiewicz logics, which though coming a lot closer, do not validate all of the laws listed. (If one defines restricted quantification in them in the obvious way, as $\forall x(Ax \rightarrow Bx)$, then they fail to validate **(I)** and **($\mathbf{I}^{-}$)** and **($\mathbf{I}_c$)**, **(III)**, **($\mathbf{III}_{var}$)**, **($\mathbf{IV}_{var}$)**, and **(VI)**, though (as we'll see in Section 12) they do obey laws which might *perhaps* be deemed "close enough approximations".) In any case, I've argued elsewhere (2003 and 2004) that the Lukasiewicz logics are simply inadequate for vagueness, for two reasons, the most serious of which is that they enable one to define a bivalent "super-determinateness" operator that obviates their whole point. The logic to be advocated below avoids both problems, as well as handling restricted quantification much better. So it's one that recommends itself for vagueness independently of issues about the semantic paradoxes; the fact that it handles the paradoxes too is an added plus for those who think a unified account is a good thing. But I will say little more about vagueness in this paper.

[1] The theory that results does share many features with the theory of Field (2008). For instance, though I won't discuss this here, we can define a determinateness operator, and iterate it transfinitely, in the same way as in that book, and the resolution of the "paradoxes of determinate truth" would carry over without significant change.

The motivation for trying to keep the above laws (together with other natural laws of restricted quantification that will emerge) isn't just that they "sound right". To those of us who are used to thinking in classical contexts, many things initially sound right that, once we take truth or vagueness seriously, are suspect except in restricted form. The motivation, rather, is that it is important to be able to reason with restricted quantifiers, and without laws that at the very least approximate these, our ability to so reason would be crippled. In Part **3** I will develop the logic of restricted quantifiers further than is required just to validate the laws above, in an effort to give initial plausibility to the idea that the construction yields a logic in which we could actually reason.

I by no means claim that the resulting logic is the best possible. Indeed, if one takes a "best possible" logic to be one whose connectives faithfully represent the connectives of natural language, the logic I will advocate is ruled out from the start. For though its conditional '$A \rightarrow B$' differs from the "material" conditional '$\neg A \vee B$' (as it must for '$A \rightarrow A$' to come out valid when '$\neg A \vee A$' isn't), still I've deliberately designed the '$\rightarrow$' so that '$A \rightarrow B$' will reduce to '$\neg A \vee B$' *in contexts where its antecedent and consequent behave classically*. That guarantees that my logic will be subject to the oddities that the material conditional has in classical contexts: for instance, it will legitimize such bizarre-sounding claims as 'If I run for President in 2016, I'll win'. My account, then, is based on the fiction that the material account of the conditional is fine for conditionals whose antecedent and consequent behave classically. The reason I've based my account on this fiction is that I don't know the nonfiction: I know of no adequate account of how the 'if ... then' of English does behave, even when its antecedent and consequent are classical. If I had such an account, I'd probably have tried to generalize that *ordinary* conditional rather than the material conditional to nonclassical contexts, and then tried to show how to combine *that* generalized conditional with naive truth and restricted quantification. (Or perhaps I'd have tried to generalize *both*: the material conditional, after all, is what is of primary use within mathematics.) But I hope and expect that the methods used for combining naive truth and restricted quantification with a nonclassical generalization of the material conditional will prove helpful to the future project of combining naive truth and restricted quantification with a nonclassical generalization of the conditional of natural language. The aim of this paper is to make progress, not to provide the ultimately best possible theory.

**§2. Naive truth and satisfaction.** By a naive truth theory, I mean a theory containing a truth predicate for its own sentences, in which a sentence attributing truth to a sentence in the language of the theory (that is unambiguous and free of indexical elements) is equivalent to the sentence itself, in the sense of being fully intersubstitutable with it (outside of quotation marks and other opaque constructions).

More exactly: when $A$ and $B$ are sentences, call $A$ **equivalent to** $B$ if, for any $X$ and $Y$ such that $Y$ results from $X$ by replacing an occurrence of $A$ with $B$ or vice versa (outside of opaque contexts), then $Y$ is inferable from $X$ and vice versa. (Obviously this suffices for multiple substitutions.) The **naivety** condition is that $True(\langle A\rangle)$ is equivalent to $A$ in this sense, where $\langle A\rangle$ is a standard name of $A$, taken to denote $A$ "as a matter of logic".[2] For instance, in such a theory

[2] If $t$ is a complex term denoting $A$ "contingently", such as "the last sentence Joe uttered", then of course $True(t)$ won't be equivalent to $A$ in the sense above. Still, either of $X$ and $Y$ will be validly inferable from the other *together with the assumption that the last sentence Joe uttered is $A$*.

(i) If $\langle A \rangle$ and $\langle B \rangle$ are true then $\langle C \rangle$ is true
will be inferable from
(ii) If $A$ and $B$ then $C$,
and vice versa, when $A$ , $B$ and $C$ are (unambiguous etc.) sentences in the language of the theory. Informally, the idea is that it would be a mistake to believe (i) but not (ii), or vice versa.

Actually I require a bit more of a naive truth theory:

(1) I require that it contain a naive satisfaction predicate for its own formulas. That's a 2-place predicate meeting the following strengthened naivety condition: that for any formula $A$ in the language with at most one free variable $x$, Satisfies($o$, $\langle A(x) \rangle$) is intersubstitutable with $A(x/o)$. (I assume that the language contains an adequate means of coding finite sequences of objects with single objects (and then decoding them), in which case a satisfaction predicate for 1-place formulas suffices for a satisfaction predicate for formulas with arbitrarily many free variables.) From such a naive satisfaction predicate, a naive truth predicate is definable.[3]

(2) I require that the theory validate some obvious composition principles, for example, that a conjunction is true if and only if both conjuncts are true, or the analog of this for satisfaction. (Such principles are generalizations whose instances are already guaranteed by what was said above; but delivering the generalizations goes beyond delivering the instances.)

(3) I require that certain schemas from the mathematics of the ground language hold even for instances containing new vocabulary such as 'True' and 'Satisfies'. I won't try here to say in general which such schemas should extend, but one that should is the induction rule $A(0), \forall n[A(n) \supset A(n+1)] \models \forall n A(n)$. (Similarly when '$\supset$' is replaced by any of the other conditionals I'll employ, such as '$\rightarrow$'.)

But the challenge to providing an adequate naive theory depends little on these extra requirements. By and large, once one has a reasonable theory with a naive truth predicate, it's easy to extend it to include a naive satisfaction predicate, and to include the required composition principles, and to validate an extended induction rule. That will be so for the construction in this paper.

Because of the possibility of "paradoxical sentences", for example, sentences asserting their own untruth, a naive truth theory is only possible in nonclassical logics. In particular, the logic in question (if it meets very minimal conditions)[4] must be either

**paraconsistent:** it must allow *acceptance* of some sentences of form $A \wedge \neg A$;[5] or
**paracomplete:** it must allow *rejection* of some sentences of form $A \vee \neg A$.

---

3 Naive theories of satisfaction are effectively equivalent to naive theories of properties: theories that for every $k$-place formula and $k$–1 objects posit a property $\lambda x A(x; o_1, ..., o_{k-1})$, such that "o instantiates $\lambda x A(x; o_1, ..., o_{k-1})$" is intersubstitutable with "$A(o; o_1, ..., o_{k-1})$". The problem discussed in this paper could equally be viewed as a problem of how to deal with restricted quantification in a naive theory of properties.

4 Basically that it include substitutivity of identity and the standard structural rules, including the transitivity of consequence. I will be tacitly assuming these things about the logic throughout.

5 This is a somewhat narrow reading of 'paraconsistent', since some relevance logicians have advocated paraconsistent logics in connection with a notion of consequence not directly connected with constraints on acceptance and rejection; but it is this narrower version of paraconsistency ("paraconsistent dialetheism") that is relevant to naive truth.

(Rejection is to be taken as a state that precludes acceptance.) Paraconsistent logics with reasonable laws of conjunction will admit **acceptance gluts**, that is, they allow the simultaneous acceptance of both a sentence and its negation. So in these logics, accepting the negation of a sentence does not require rejecting the sentence, given that rejection precludes acceptance. Dually, paracomplete logics with reasonable laws of disjunction will admit **rejection gluts**, that is, the simultaneous rejection of both a sentence and its negation. So in them, rejecting a sentence does not require accepting its negation. (In particular, rejection of some instances of excluded middle doesn't require accepting the negations of those instances.)[6] The logic could be both paraconsistent and paracomplete, either (i) by allowing accepting both $A$ and $\neg A$ while simultaneously rejecting both $B$ and $\neg B$ for a $B$ that differs from $A$, or (ii) by allowing for either accepting both $A$ and $\neg A$ or for rejecting them both (though of course, not doing both at once).

There should be no worry that the use of a nonclassical logic will upset standard mathematics, because we will be able to confine the nonclassicality so that it doesn't arise within standard mathematics. (We can use the truth or satisfaction predicate to *extend* standard mathematics, for example, by extending the induction rule; the *extension* will involve some nonclassicality, as it must to maintain naivety while avoiding paradox. But the unextended mathematics will be entirely classical, as indeed will large parts of the extension: each part of the standard Tarski hierarchy will be a classical fragment of the extended mathematics.)

The logic I will be suggesting in this paper is paracomplete but not paraconsistent: it permits rejection gluts but not acceptance gluts. As I'll note in the final section, the goal of maintaining the above menu of laws of restricted quantification is not achievable in a paraconsistent naive theory, and I'm skeptical that one can come satisfactorily close.

I take it that the idea of paracomplete logic for dealing with the semantic paradoxes is highly natural. Let $Q$ abbreviate a Liar sentence, that is, a sentence asserting its own untruth, that is, a sentence equivalent to the claim $\neg True(\langle Q \rangle)$ where $\langle Q \rangle$ is a term denoting $Q$. It's natural to think that it would be absurd to assert

---

[6] Both sorts of logics allow both acceptance *gaps* and rejection *gaps*, but that has nothing special to do with these logics: any reasonable logic will allow for the nonacceptance of both 'This coin will land heads' and its negation, and also for the nonrejection of both.

Many have tried to characterize paraconsistent and paracomplete approaches by saying that paraconsistent logics are those that allow for the acceptance of *truth value gluts* (sentences that are both true and false, where falsity is truth of negation) and that paracomplete logics are those that allow for the acceptance of *truth value gaps* (sentences that are neither true nor false). This misfires for a number of reasons.

(1) There are well-known truth theories in classical logic (hence not paracomplete) that accept truth value gaps; and duals to these, also in classical logic (and hence not paraconsistent), that accept truth value gluts. (Of course, neither of these can be naive truth theories, since no classical theory can be naive.)

(2) In typical naive truth theories (those where $\neg\neg A$ is equivalent to $A$), *there is no distinction between truth value gaps and truth value gluts:* the "gap" statement $\neg True(\langle A \rangle) \wedge \neg True(\langle \neg A \rangle)$ is equivalent to $\neg A \wedge \neg\neg A$, hence to $\neg A \wedge A$, hence to the "glut" statement $True(\langle A \rangle) \wedge True(\langle \neg A \rangle)$. There is only a notion of a "gap/glut", which is as much one as the other.

Naive paracomplete theories (unless they are paraconsistent as well) do not accept truth value gaps (equivalently, truth value gluts), indeed they reject them. It is the issue of whether to accept gap/gluts or to reject them that differentiates the paraconsistent and paracomplete approaches. To put the distinction in terms of the acceptance of truth value gluts versus the acceptance of truth value gaps is totally misguided.

(i) $Q \wedge \neg True(\langle Q \rangle)$:

the second conjunct seems to undermine the first. Similarly, it's natural to think that it's absurd to assert

(ii) $True(\langle Q \rangle) \wedge \neg Q$ :

again, the second conjunct seems to undermine the first. But in (i) the two conjuncts are equivalent; similarly in (ii), assuming the redundancy of double negation. Given this, it seems that each of

(i*) $\neg True(\langle Q \rangle)$
(ii*) $True(\langle Q \rangle)$

must be absurd, so we should fully reject them. This gives us a rejection glut. And with the further natural view, that if we fully reject each of two claims we should fully reject their disjunction, we get paracompleteness in the form I officially defined it. (Of course, committed proponents of alternative treatments of the paradoxes will reject some of the assumptions I've called "natural", but my goal is only to provide initial motivation for the approach I'll be taking.)

It's well-known that no logic (meeting the assumptions of note 4) can accord with naive truth if it has a conditional that both obeys modus ponens and allows unrestricted conditional proof: that's the lesson of Curry's paradox. Since I'm keeping modus ponens without restriction, there will need to be restrictions on conditional proof; this will allow there to be valid rules $A \vDash B$ for which the corresponding conditional is not valid. There will however be useful restricted versions of conditional proof; indeed, the natural way to establish the laws mentioned in Section I (which are all valid conditionals, not merely valid rules) will be by restricted conditional proof.

**§3. The restricted quantifier and the conditional.** I'll use the notation $\forall x(Ax/Bx)$ for "All $A$ are $B$". But rather than taking this as a primitive binary quantifier, I'll take it to be defined from the ordinary universal quantifier and a conditional $\blacktriangleright$: $\forall x(Ax/Bx)$ will abbreviate $\forall x(Ax \blacktriangleright Bx)$. That this can be done isn't a substantive assumption: one could equally well take the restricted quantifier as primitive and define $A \blacktriangleright B$ as $\forall x(A/B)$ where $x$ is a variable not occurring in either $A$ or $B$. (Or if one doesn't like quantification by a variable not free in the formulas to which the quantifier is attached, define it as $\forall x(x = x \wedge A\, /x = x \wedge B)$.) With only very minimal assumptions on how the conditional and the restricted quantifier behave, it's easy to see that whichever of the two one starts with, if one defines the other in terms of it as above and then does the reverse definition, then one gets something equivalent to what one started from. I prefer defining the restricted quantifier in terms of the conditional only because it makes the development to follow more perspicuous.

What *would* be a substantive assumption is that the conditional that's interdefinable with universal restricted quantification is the same conditional as the 'if ... then' that appears in proposed the laws of Section I.[7] In order to avoid such an assumption, let's use a different

[7] For what it's worth, those who think of the semantics of 'if...then' as given in part by Adams' Thesis (hence differing from $\supset$) are likely to think that in classical contexts restricted quantification is to be explained by the material conditional, that is, that $\blacktriangleright$ is just $\supset$. So they will share the view that $\blacktriangleright$ differs from 'if...then'. This point may not be *directly* relevant to the present

notation for the latter; I'll use $\rightarrow$. It will turn out that there are reasons for taking the two conditionals to be distinct from each other, and also to be distinct from the conditional $\supset$ defined by $\neg A \vee B$. Nonetheless, we'll see that both these new conditionals are equivalent to $\supset$ "in classical contexts", on a suitable explication of that phrase. As a very special case, when neither $A$ nor $B$ contains 'True' (and we don't allow for vagueness), any one of $A \rightarrow B$, $A \blacktriangleright B$ and $A \supset B$ will be validly intersubstitutable with any other.

Given the definition of $\forall x(Ax/Bx)$ in terms of $\blacktriangleright$, it's clear what's required of $\rightarrow$ and $\blacktriangleright$ for the laws of Section 1 to be valid; it's that all instances of the following schemas be valid:

**(I-$\blacktriangleright$):** $[(A \blacktriangleright B) \wedge A] \rightarrow B$

**(II-$\blacktriangleright$):** $B \rightarrow (A \blacktriangleright B)$

**(III-$\blacktriangleright$):** $[(A \blacktriangleright B) \wedge (B \blacktriangleright C)] \rightarrow (A \blacktriangleright C)$

**(III-$\blacktriangleright_{\text{var}}$):** $[(A \blacktriangleright B) \wedge \neg(A \blacktriangleright C)] \rightarrow \neg(B \blacktriangleright C)$

**(IV-$\blacktriangleright$):** $[(A \blacktriangleright B) \wedge (A \blacktriangleright C)] \rightarrow [A \blacktriangleright (B \wedge C)]$

**(IV-$\blacktriangleright_{\text{var}}$):** $[[(A \blacktriangleright B) \wedge \neg[A \blacktriangleright (B \wedge C)]] \rightarrow \neg(A \blacktriangleright C)]$

**(V-$\blacktriangleright$):** $\neg(A \blacktriangleright B) \rightarrow (A \wedge \neg B)$

**(VI-$\blacktriangleright$):** $(A \wedge \neg B) \rightarrow \neg(A \blacktriangleright B)$

**(C-$\blacktriangleright$):** $(A \blacktriangleright B) \leftrightarrow (\neg B \blacktriangleright \neg A)$, where $C \leftrightarrow D$ abbreviates $(C \rightarrow D) \wedge (D \rightarrow C)$.

A look at (I-$\blacktriangleright$) makes clear why I wanted two separate conditionals: no naive truth theory satisfying my background conditions can have the analog of (I-$\blacktriangleright$) with both conditionals the same. (For any conditional $\gg$, the alleged law $[(A \gg B) \wedge A] \gg B$ is called **pseudo modus ponens** for $\gg$, and it is well known that we can't have both it and genuine modus ponens for $\gg$ (the rule of inference from $A \wedge (A \gg B)$ to $B$) without triviality, in a naive theory satisfying the conditions of note 4.) In order to have a chance of obtaining (I-$\blacktriangleright$) (in conjunction with modus ponens for $\blacktriangleright$), the inference from $A \rightarrow B$ to $A \blacktriangleright B$ must fail.

A look at the laws also reveals that unlike $\blacktriangleright$, we better not take $\rightarrow$ to be in general contraposable: we better not accept the inference from $A \rightarrow B$ to $\neg B \rightarrow \neg A$ in general, though it will be fine when $A$ is classical.[8] Here, in sketch. are three routes to the conclusion that the general contraposability of $\rightarrow$ would cause difficulties:

(1) (Perhaps the most obvious, using (VI-$\blacktriangleright$):) Were it contraposable, one could get from (VI-$\blacktriangleright$) to $(A \blacktriangleright B) \rightarrow \neg(A \wedge \neg B)$, and in particular $(A \blacktriangleright A) \rightarrow \neg(A \wedge \neg A)$; hence assuming $A \blacktriangleright A$ and modus ponens for $\rightarrow$, we'd have $\neg(A \wedge \neg A)$. But that is equivalent to excluded middle if we assume a deMorgan's law and the redundancy of double negation; so if we assume the latter laws, as I will, a logic without excluded middle that keeps (VI-$\blacktriangleright$) can't allow $\rightarrow$ to contrapose.

project, given that that project assumes an 'if...then' that reduces to the material conditional in classical contexts (see the end of Section 1); still, it might make the suggestion of an $\blacktriangleright$ distinct from 'if ... then' seem less *ad hoc*.

[8] It's perhaps worth noting that if we were engaged in giving a theory of the '$\rightarrow$' that faithfully reflected the 'if...then' of English, we wouldn't want 'if ...then' to contrapose even in contexts where the antecedent and consequent are classical. (The noncontraposability of the English conditional is well-known: for example, from 'If Clinton won't be the next President, Bachmann won't either' I can't infer 'If Bachmann will be the next President, Clinton will too'.) As with the point made in the previous footnote, this is only indirectly relevant in the present context, but it doesn't seem entirely irrelevant.

(2) (Using (I-$\blacktriangleright$) and others instead of (VI-$\blacktriangleright$):) (II-$\blacktriangleright$) and (C-$\blacktriangleright$) lead to $\neg A \rightarrow (A \blacktriangleright B)$, but this and (I-$\blacktriangleright$) will lead to $[A \wedge \neg A] \rightarrow B$ by laws we will surely want, which by contraposition of $\rightarrow$ would yield $\neg B \rightarrow \neg(A \wedge \neg A)$. So if $B$ is absurd we'd get $\neg(A \wedge \neg A)$ by modus ponens, which again leads to excluded middle by deMorgan's and double negation.

(3) (Using (III-$\blacktriangleright$) and others, but not (I-$\blacktriangleright$) and only the rule form of (VI-$\blacktriangleright$):) Let $\top$ be a logical truth. The rule form of (VI-$\blacktriangleright$) yields $\neg(\top \blacktriangleright \neg\top)$. In the contraposed form of (III-$\blacktriangleright$), let $A$ be $\top$ and $C$ be $\neg\top$; then applying modus ponens we get $\neg[(\top \blacktriangleright B) \wedge (B \blacktriangleright \neg\top)]$; or using deMorgan's, $\neg(\top \blacktriangleright B) \vee \neg(B \blacktriangleright \neg\top)$. Using (V-$\blacktriangleright$) twice together with obvious laws of disjunction we get $\neg B \vee B$.

(All the routes seem to require the deMorgan's law $\neg(A \wedge B) \models \neg A \vee \neg B$, so I suppose one might try keeping contraposability and rejecting that; it isn't an option I've explored.)

Once we realize that the 'if ... then' may not be contraposable, the question arises as to which of the contrapositions of principles on our original list ought to hold. I think that there is a strong intuitive case for wanting the following additions to our list:

**(II*):** If not all $A$ are $B$, then not everything is $B$.

**(IV*):** If not all $A$ are both $B$ and $C$, then either not all $A$ are $B$ or not all $A$ are $C$.

**(V*):** If everything is either not $A$ or $B$, then all $A$ are $B$. (That is, if nothing is both $A$ and not $B$ then all $A$ are $B$.)[9]

We probably also want

**(C*):** Not all $A$ are $B$ if and only if not all not-$B$ are not-$A$

Adding all these to our original list amounts to strengthening (II-$\blacktriangleright$), (IV-$\blacktriangleright$) (V-$\blacktriangleright$) and (C-$\blacktriangleright$) by replacing $\rightarrow$ with $\rightarrow\rightarrow$ and $\leftrightarrow$ with $\leftrightarrow\leftrightarrow$, where $A \rightarrow\rightarrow B$ abbreviates $(A \rightarrow B) \wedge (\neg B \rightarrow \neg A)$ and $A \leftrightarrow\leftrightarrow B$ abbreviates $(A \rightarrow\rightarrow B) \wedge (B \rightarrow\rightarrow A)$. (In fact, I'm inclined to think that even **(C)** and (**C***) together understate the intuitive requirements: we should have that "All $A$ are $B$" and "All not $B$ are not $A$" should be fully equivalent, that is, intersubstitutable in all contexts, even inside both sorts of conditionals.) My account will deliver all these things.[10]

Although I prefer to define restricted quantification in terms of a contraposable $\blacktriangleright$, it is convenient to define that in turn from a noncontraposable conditional $\rhd$: $A \blacktriangleright B$ will be

---

[9] Note that **(V*)** entails **(II$_c$)**. Similarly, the old **(V)** entails the following analog of **(II$_c$)**:

**(II*$_c$)** If not all $A$ are $B$, then something is not $B$.

[10] One has two choices for a restricted existential quantifier "Some $A$ are $B$": either $\neg\forall x(Ax/\neg Bx)$ or $\exists x(Ax \wedge Bx)$. Laws **(V)** and **(V*)** will be seen to entail that the first of these implies the second outside of $\blacktriangleright$ contexts. I find the second ("weaker") reading of "Some $A$ are $B$" far more natural, but to avoid any confusion, I'll avoid the locution "Some $A$ are $B$" in this paper.
There are also two choices for reading "No $A$ are $B$", namely $\forall x(Ax/\neg Bx)$ or $\neg\exists x(Ax \wedge Bx)$. In the case where $Ax$ is "$x = x \wedge Q$" (where $Q$ is the Liar sentence), the first makes "No $A$ are $A$" and "No $A$ are not $A$" fully acceptable; the latter doesn't, but doesn't make their negations acceptable either. I'm inclined to go for the latter reading here too, making "No $A$ are $B$" the negation of "Some $A$ are $B$" rather than the equivalent of "All $A$ are not $B$"; but I'll also avoid the locution "No $A$ are $B$" in what follows.

$(A \rhd B) \wedge (\neg B \rhd \neg A)$, making its contraposability evident.[11] But the main focus will be on $\blacktriangleright$ rather than $\rhd$.

How will the contraposable $\blacktriangleright$ connect to $\supset$? Letting $\vDash$ mean validity (which won't be officially defined until Section 9, but I assume the reader to have a pretheoretic grasp of it), we will have the following:

**Limited $\blacktriangleright$ collapse:**

(i) $\vDash (A \supset B) \rightarrow (A \blacktriangleright B)$

(ii) $\vDash \neg(A \blacktriangleright B) \rightarrow \neg(A \supset B)$

(iii) $\vDash \neg(A \supset B) \rightarrow \neg(A \blacktriangleright B)$

(iv) $\top \rightarrow [(A \vee \neg A) \vee (B \vee \neg B)] \vDash (A \blacktriangleright B) \rightarrow (A \supset B)$

Since $\rightarrow$ will obey modus ponens, we will also get weaker $\rightarrow$-free forms of all these, for example,

($\mathrm{i_W}$) $A \supset B \vDash A \blacktriangleright B$.

The premise of (iv) (and its $\rightarrow$-free weakening) is a slight strengthening of the claim that excluded middle holds *either* for $A$ *or* for $B$. Such a premise is required, since we will want for instance $Q \blacktriangleright Q$ but not $Q \supset Q$ when $Q$ is a Liar sentence. (We *don't* need excluded middle for *both* $A$ and $B$.)

It might seem tempting to summarize this limited collapse result by saying that $A \blacktriangleright B$ is weaker than $A \supset B$, but equivalent to it on the strengthened excluded middle assumption. However, such a summary would be unfortunate, for the limited result tells us nothing about how the embeddings of $A \blacktriangleright B$ and $A \supset B$ inside the conditionals $\rightarrow$ and $\blacktriangleright$ (or $\rhd$) relate to each other. The equivalence of $A$ and $B$, as defined in Section 2 in terms of intersubstitutivity, requires more than the validity of the biconditionals $A \leftrightarrow B$ and $\neg A \leftrightarrow \neg B$. Rather, equivalence as there defined relates to the validity of a different biconditional $\Leftrightarrow$, defined from $\rightarrow$ and $\blacktriangleright$ together. Let $A \Rightarrow B$ be short for $\top \rightarrow (A \blacktriangleright B)$, where $\top$ is any logical truth (e.g., $\forall x(x = x)$). $\Rightarrow$ will be contraposable.[12] Let $\Leftrightarrow$ be the corresponding biconditional, that is, $(A \Rightarrow B) \wedge (B \Rightarrow A)$; then it will turn out that $A$ is equivalent to $B$ in the sense of valid intersubstitutability if and only if $\vDash A \Leftrightarrow B$.

Given this, it's natural to say that $A$ is **at least as strong as** $B$ if $\vDash A \Rightarrow B$; by contraposability, $A$ is at least as strong as $B$ only if $\neg B$ is at least as strong as $\neg A$. Also, we can call $A$ **logically comparable to** $B$ if either $A$ is at least as strong as $B$ or $B$ is at least as strong as $A$. On this definition, $A \blacktriangleright B$ will turn out in general to be incomparable to $A \supset B$. Nonetheless, $A \blacktriangleright B$ will reduce to $A \supset B$ in suitably classical contexts: that is, they will turn out to be equivalent given the assumption of a strong form of excluded middle for *both* $A$ and $B$. (Indeed, $A \rhd B$ will reduce to $A \supset B$ as well.)

We wouldn't expect a "Limited $\rightarrow$-collapse" result analogous to the above, because of the noncontraposability of $\rightarrow$. (But as mentioned, we do get a collapse result for $\rightarrow$: it,

[11] One reason for doing this is to allow anyone who doesn't think restricted quantification should contrapose to define it in terms of $\rhd$ rather than $\blacktriangleright$. In addition, it is natural to investigate how our two conditionals $\rightarrow$ and $\blacktriangleright$ compare, and it is easier to do this if we first strip off the differences that result from the contraposability of the latter. Finally, there is a connective $\Rightarrow_{nc}$ that can be defined from $\rightarrow$ and $\rhd$ together, that is useful even for exploring the properties of $\blacktriangleright$.

[12] The conditional $A \Rightarrow_{nc} B$, mentioned in the previous footnote, will be defined analogously as $\top \rightarrow (A \rhd B)$, and is noncontraposable.

like $\blacktriangleright$ and $\rhd$, will reduce to $\supset$ in classical contexts.) Still, we will get an analog of (ii) (and hence ($\mathrm{ii_W}$)) and of ($\mathrm{iii_W}$):

$\models \neg(A \rightarrow B) \rightarrow \neg(A \supset B)$ and

$\neg(A \supset B) \models \neg(A \rightarrow B)$.

Using these and the corresponding results for $\blacktriangleright$, we get

$\models \neg(A \rightarrow B) \rightarrow \neg(A \blacktriangleright B)$ and $\neg(A \blacktriangleright B) \models \neg(A \rightarrow B)$

(and the analogs of these with $\rhd$ in place of $\blacktriangleright$). We clearly *cannot* have $A \rightarrow B \models A \blacktriangleright B$; that with (I-$\blacktriangleright$) would yield pseudo-modus-ponens for $\blacktriangleright$. As it happens, we won't have even $A \rightarrow B \models A \rhd B$, and we won't have $A \blacktriangleright B \models A \rightarrow B$ either.

More on these matters in due course.

**§4. Validity, conservativeness, and general strategy.** My concern is with getting a system of logic in which a great many laws involving truth, satisfaction, and the two conditionals are valid (i.e., have only valid instances; laws themselves I take as general schemas, showing valid logical forms).[13] A proper model-theoretic account of validity involves quantifying over models (of an appropriate sort) for a language that includes these notions: the inference from $\Gamma$ to $B$ is valid if *for all models of such and such kind*, .... It's important for any such definition that the class of models be fairly wide in order not to declare valid things that really hold only in the particular models looked at.

Some restrictions on the language are inevitable, since in order to sensibly talk of truth we need to presuppose a background theory of the bearers of truth, viz. sentences; and in order to sensibly talk of the satisfaction of formulas (with arbitrarily many free variables) by finite sequences, we must presuppose a background theory not only of formulas but of arbitrary finite sequences of the objects that the theory recognizes. Presumably this background theory is to be stated in the **ground fragment** of the language $L$, that is, the maximal fragment $L_0$ not containing 'True' or 'Satisfies' or the new conditionals. So in giving a model-theoretic account of validity for a language with 'True' or 'Satisfies' (and maybe the new conditionals), we must restrict attention to languages in whose ground fragments we can state the required background theory; I'll call such languages **expressively adequate**. Moreover, we must restrict attention to models whose ground fragments satisfy this background theory. It is common to take the background theory to explicitly include arithmetic, and to identify expressions of the language with numbers by a Gödel coding, though of course this is an inessential matter of convenience. I'll also assume that the models in question are **$\omega$-standard**: when the background theory includes arithmetic, this means that the arithmetic part of the model is just the standard model of arithmetic. One rationale for requiring $\omega$-standardness is that it guarantees that once we extend the model to cover a truth or satisfaction predicate, the rule of mathematical induction will apply even to formulas containing the new truth or satisfaction predicate.[14] By a **minimally acceptable model** of an expressively adequate language I'll mean an $\omega$-standard model of it that satisfies the aforementioned background theory. (This is a requirement on the ground fragment only.)

---

[13] I will use upper case letters like $A$ , $B$ etc. ambiguously, both as schematic letters taking formulas as instances and as meta-linguistic variables over formulas. (Indeed, I've already done so.) The danger of this leading to confusion is almost nonexistent.

[14] Another rationale for requiring $\omega$-standardness isn't for arithmetic *per se*, it's for syntax: it's to ensure that for each $o_1$ in the domain of the model, there can only be finitely many $o_2$ in the domain for which $< o_1, o_2 >$ satisfies '$x_1$ and $x_2$ are expressions of $L$ and $x_2$ is part of $x_1$'.

To avoid notational complexity I'm going to restrict attention in the paper to languages that contain truth predicates but not satisfaction predicates. (So the languages won't be able to express the compositional laws for quantifiers.) This would be a serious limitation if there were any reason for the restriction beyond convenience and readability; but in fact it is routine if notationally messy to extend to satisfaction (and the compositional laws will then hold).

Here's an impressionistic sketch of the strategy to follow; it will be made more precise as I go on.

Let $L$ be any expressively adequate language with 'True', '$\triangleright$' and '$\rightarrow$'. Part **2** of the paper will be concerned with showing how any minimally acceptable classical model $M$ for its ground fragment can in a certain sense be "extended" to a multi-valued model $\Psi(M)$ for all of $L$ which (in a sense to be explained) validates the desired laws. The "extension" does not involve an extension of the domain: $\Psi(M)$ will have exactly the same domain as $M$. Moreover, individual constants will name the same things in $\Psi(M)$ as they do in $M$, and function symbols will stand for the same operations in $\Psi(M)$ as in $M$. Furthermore, predicates in the ground language will behave in the same way in $\Psi(M)$ as in $M$ (though stating this carefully requires a correspondence between the two classical semantic values 0 and 1 of $M$ and two of the values $\mathbf{0}_M$ and $\mathbf{1}_M$ employed in $\Psi$(M)). The "extension" is simply that now sentences containing 'True', '$\triangleright$' and '$\rightarrow$' get values—in many cases nonclassical values, that is, values not corresponding to the two classical values.

An important feature of this procedure is

**(CNSV):** For any sentence of the ground language, its value in $\Psi(M)$ is $\mathbf{1}_M$ if its value in $M$ is 1, and its value in $\Psi(M)$ is $\mathbf{0}_M$ if its value in $M$ is 0.

Call a set $\Gamma$ of sentences of the ground language ***-consistent** if there is a minimally acceptable classical model in which all its members are true. Since every model of form $\Psi(M)$ will "validate" the desired laws (in the sense to be explained), and will also "validate" any sentence that gets value $\mathbf{1}_M$, it follows from (CNSV) that for every *-consistent set $\Gamma$ of sentences in the ground language, there is at least one nonclassical model which validates all its members and also validates the desired laws. Or as we could naturally put it:

**(CONSERV):** Any classically *-consistent set of sentences in the ground language is *-consistent in the desired logic.

I take that to be a desideratum of an adequate theory.[15]

(We will actually be able to extend (CONSERV) in a useful way. For we will be able to define in the logic a predicate 'classical' such that when you add to any classical argument the additional premise that all the formulas in it are classical, the resulting argument is valid in the new logic.[16] Then the strengthened result is that if $\Gamma$ is a classically *-consistent set

[15] Does this conflict with well-known results on the "nonconservativeness of truth" (reviewed in Halbach, 2011 and Horsten, 2011)? The compositional laws will hold in the logic (or at least, they would if the construction were done with a satisfaction predicate, to allow the composition laws for the quantifiers to be expressed), and I've already noted that the rule of mathematical induction will extend to formulas containing 'True' and 'Satisfies'; in these circumstances, there's a sense in which the addition of 'True' or 'Satisfies' can't be a conservative extension. But what this means is simply that *the analog of (CONSERV) with consistency instead of *-consistency* can't hold. The nonconservativeness results depend on including models that are not $\omega$-standard, and by moving to *-consistency I've excluded them.

[16] 'Classical' must be defined using 'True', and excluded middle does not apply to it; this is inevitable, for otherwise we could get a Liar-like paradox, from a sentence that asserts of itself that it is not both classical and true.

of sentences in the ground language, then Γ *augmented with the claim (CLASS) that all sentences in the ground language are classical* is *-consistent in the new logic. So we can safely assume (CLASS) in the new logic, and when we do, all classical proofs in the ground language carry over to the new logic essentially unchanged.)

It's clear that to carry out the strategy I've outlined, we must run the construction in such a way that it works for every minimally acceptable ground model. I will do so.[17] Since the construction will not depend on the details of the ground model, for much of the paper we will be able to hold that model fixed; but it is important to remember that it is really a construction for any ground model that meets the minimal presuppositions required for the notions employed (such as truth) to even make sense.

## Part 2: THE CONSTRUCTION

**§5. Overview; valuations, and boosters.** Let $L$ be a language with 'True', '$\rightarrow$' and '$\triangleright$', whose ground fragment $L_0$ is expressively adequate for the syntax of $L$, and let $M$ be any minimally adequate classical model of $L_0$. (The terminology was defined early in Section 4.) The goal is to construct some sort of multi-valued model $\Psi(M)$ based on $M$, in which sentences in the larger language $L$ get values, and in which every sentence in the original language gets the same value it has in $M$. (I remind you that I'm sticking to 'True' rather than the more general 'Satisfies' only for notational convenience: the construction generalizes in a straightforward manner.)

To be more explicit, I'll take a **classical model** $M$ for the ground fragment $L_0$ to consist of

- a nonempty domain;
- an assignment of an object in that domain to each individual constant, and of an $n$-ary operation on that domain to each $n$-ary function symbol;
- an assignment of an **$n$-ary 2-extension** to each $n$-place predicate *of the ground fragment*, where by an $n$-ary 2-extension I mean an $n$-place function from the domain to the set $\{0,1\}$. (For classical models, value 1 corresponds to truth in the model, 0 to falsity in the model.) I'll also require that the 2-extension of '=', which presumably is in the language, be the function that maps each pair of form <o,o> into 1 and each other pair into 0.

An **assignment function** for a ground model $M$ will be a function that assigns to each variable of the language an object in the domain of $M$.

We can think of the $\Psi(M)$ to be constructed as based on an infinite set $\mathbf{F}_M$ of "3-valued worlds", all of which are just like $M$ as far as ground level formulas are concerned. (In particular, all have the same domain.) In calling a member of $\mathbf{F}_M$ a "3-valued world", I mean that at it, each formula of the full language $L$ is to get one of three values, relative to $M$ and to any assignment function $s$. (Despite the "worlds" terminology,

(On another matter, it makes no difference whether 'all formulas in it are classical' is formulated with a $\supset$ or a $\blacktriangleright$, given Limited $\blacktriangleright$-collapse and the fact that '$x$ is a formula in argument $a$' is bivalent.)

[17] Acceptable ground models for standard set theories like ZFC are of special interest, since the construction in this paper can be carried out in them; nonetheless it would be inappropriate to consider *only* such models, since that would blur the distinction between mathematics and logic by making every sentence of the set theory come out valid.

there isn't any real connection to necessity or possibility, just a formal analogy to modal semantics.)

An important remark about the three values: I will call them 1, 0, and $\frac{1}{2}$, but you are *not* to think of them as meaning 'true (at the world)', 'false (at the world)' and 'neither (at the world)'. That can only lead to confusion. It's only slightly better to make the model-relativity explicit, and speak of 'true at the world, *in the model*', etc.. Rather, the claim that a sentence has value $\frac{1}{2}$ at a world of the model says nothing one way or the other about whether it is "true in the model at that world": all we can say is that the claim that it is true has value $\frac{1}{2}$ at the world in the model. Any attempt to talk in a bivalent way about what's true at that world in the model will lead to all sorts of problems.

$\Psi(M)$ itself is the product of the 3-valued worlds in $\mathbf{F}_M$: that is, in $\Psi(M)$ the value of a formula $A$, relative to assignment function $s$, is the function that assigns to each world in $\mathbf{F}_M$ the value in $\{0,\frac{1}{2}, 1\}$ that it gets in that world. $\Psi(M)$ thus assigns values $f$ in the space $3^{F_M}$ (or a sufficiently inclusive subset $\mathbf{V}_M$ of $3^{F_M}$). The values are partially ordered by the relation $f \preceq g$, defined as $(\forall z \in \mathbf{F}_M)(f(z) \leq g(z))$; there is a maximum value $\mathbf{1}_M$ (the function that assigns 1 to every world in $\mathbf{F}_M$) and a minimum value $\mathbf{0}_M$ (defined analogously). A formula in the ground language that gets value 1 in $M$ will get value $\mathbf{1}_M$ in $\Psi$(M); similarly if it gets 0 in $M$ it will get $\mathbf{0}_M$ in $\Psi$(M). (When it comes to defining validity, I will use for each $M$ an upwards-closed set of designated values that includes $\mathbf{1}_M$.)[18]

Each "world" will be obtained from the classical starting model $M$ by adding three things:

- a **3-extension** $T$ for 'True', where a 3-extension is a function from $M$ to $\{0, \frac{1}{2}, 1\}$.
- a **$\rhd$-valuation** $\upsilon$, where that is a function that takes any assignment-function $s$ and any $L$-formula of form $A \rhd B$ into a member of $\{0, \frac{1}{2}, 1\}$, subject to the restriction that if the free variables of $A \rhd B$ are all in a set $Y$, any two assignment functions that are alike outside of $Y$ give the same value to $A \rhd B$.
- an **$\rightarrow$-booster** (or **$\rightarrow$-prevaluation**) $g$, defined analogously except with $A \rightarrow B$ in place of $A \rhd B$, *and which can assign only values 1 or* $\frac{1}{2}$. ($g$ thus simply serves to codify a binary relation.)

For the moment, these will be arbitrary, though we'll eventually impose strong restrictions on which ones we consider. (In Section 6 we will choose, for each pair of $g$ and $\upsilon$, a single 3-extension of 'True' to go with them. In Section 7 we will choose, for each $g$, a "fiber" $\mathbf{R}_g$ of those $\upsilon$ to consider in connection with $g$. And in Section 8 we will choose a set $\mathbf{J}$ of appropriate $g$. So the worlds we eventually get are determined by pairs $< g, \upsilon >$ where $g \in \mathbf{J}$ and $\upsilon \in \mathbf{R}_g$; for each choice of $g$ and $\upsilon$ there will be a uniquely determined $T$.)

The remaining task of this section is to give the valuation rules for arbitrary "3-valued worlds" (before the additional restrictions just mentioned are imposed). Basically these will be what's known as the Strong Kleene rules, but there will be one minor surprise in the differential treatment of $\upsilon$ and $g$ (which will explain why I'm calling one of them a 'valuation' and the other a 'prevaluation' or 'booster'.)

---

[18] Note that the space $\mathbf{F}_M$ and the value space $V_M$ based on it depend on the starting model $M$. In particular, as the cardinality of $M$ increases, that of $\mathbf{F}_M$ and $V_M$ will typically increase. Everything here is model-relative; the construction will give no sense to any model-independent notion of a sentence's value. That is an additional reason why we should not try to associate any of the values with truth.

Before turning to this, I'll make another notational simplification. I'll need to consider valuations of formulas relative to several parameters, and having additional subscripts for the assignment function would make for a very cumbersome notation. A common method for avoiding this, which I'll adopt, is to consider an expanded language $L^+_M$ (possibly with a very large cardinality of symbols) with a name for every object in $M$, in which case one can treat quantification substitutionally.[19] (I'll usually just write '$L^+$', leaving the subscript '$M$' implicit.)

This is a familiar device, but in the context of a truth theory, one must be careful to remember that the truth predicate we're investigating is the one for the unexpanded $L$, not $L^+$—$L^+$ is simply a technical artifice, and one that depends on a specific choice of $M$. (Sentences of $L^+$ needn't even be in the domain of the model; and even if we think of the domain as never really including sentences themselves but only Gödel codes for them, the Gödel coding is for $L$ rather than for $L^+$; $L^+$ might have too many names to be representable by any normal kind of Gödel coding.) So we need to carefully distinguish when we're talking about terms, formulas and sentences of $L^+$ and when we're talking about the more restricted class of terms, formulas and sentences of $L$. There are some obvious rules of thumb,[20] but I'll try to be explicit whenever there is any possibility of confusion.

Using this artifice, we can describe $\rightarrow$-boosters and $\rhd$-valuations more simply than I did above. Instead of an $\rightarrow$-booster or $\rhd$-valuation being a function that assigns values to *pairs of assignment functions and L-formulas* of conditional form, it is a function assigning values to *$L^+$-sentences* of conditional form.

And now I give the valuation rules for arbitrary 3-worlds, based on the classical model $M^+$ of the ground fragment of the expanded ground language $L^+_M$. (For readability I will leave the model $M^+$ out of the notation, and when I do include it I will just call it $M$.)

- If $p$ is a predicate of the ground language, $|p(t_1, ..., t_n)|_{g,v,T}$ is the result of applying $\|p\|$ (the 2-extension assigned to $p$) to $< den(t_1), ..., den(t_n) >$
- $|True(t)|_{g,v,T}$ is $T(den(t))$ [$T$ is as yet undetermined]
- $|A \rhd B|_{g,v,T} = v(A \rhd B)$ [$v$ is as yet undetermined; $g$ and $T$ drop out as irrelevant, as does any structure within $A \rhd B$]
- $|\neg A|_{g,v,T} = 1 - |A|_{g,v,T}$
- $|A \wedge B|_{g,v,T} = min\{|A|_{g,v,T}, |B|_{g,v,T}\}$

[19] Of course there's no loss in expressive power: for a formula $A$ (of $L$ or even $L^+$), $|A|_{g,v,s}$ is definable as $|A^*|_{g,v}$ where $A^*$ is an $L^+$−sentence that is just like $A$ but with an $L^+$-name for $s(v)$ substituted for all free occurrences of $v$, for each $v$ free in $A$.

[20] (i) When I discuss the semantic value of an $A$ relative to some parameters that don't include an assignment function, for example, by using a notation such as $|A|_{g,v}$, the $A$ involved will be assumed to be a sentence, and usually I'll allow it to be a sentence of $L^+$ rather than requiring it to be in $L$.
(ii) As a special case of (i), when I discuss the semantic value of something of form $A(t)$ relative to some parameters that don't include an assignment function, for example, by using a notation such as $|A(t)|_{g,v}$, the $t$ involved will be assumed to be a closed singular term, and normally allowed to be in $L^+$ rather than requiring it to be in $L$.
(iii) When I say such things as "If $t$ is a term that denotes $B$, then $|True(t)|_{g,v}$ is ....", $B$ will normally be assumed to be a sentence of $L$, given that we're going to want 'True' to trivialize (always have value 0) in application to anything that isn't a sentence of $L$.

- $|A \to B|_{g,v,T} = \begin{cases} 1 & \text{if } g(A \to B) = 1 \\ 0 & \text{if } g(A \to B) < 1 \text{ and } |A|_{g,v,T} = 1 \text{ and } |B|_{g,v,T} = 0 \\ \frac{1}{2} & \text{otherwise} \end{cases}$
- $|\forall x A|_{g,v,T} = min\{|A(x/t)|_{g,v,T} : t$ is a closed $L^+$-term$\}$, where $A(x/t)$ is the result of substituting $t$ for $x$.

The definitions of $\vee$ and $\exists$ in terms of $\neg$, $\wedge$ and $\forall$ yield analogs of the $\wedge$ and $\forall$ clauses, with maximum instead of minimum.

Using these rules, we determine a value $|A|_{g,v,T}$ for every sentence of the language $L^+$ on the basis of complexity. ($\rhd$-sentences can be treated as of complexity 0, though they needn't be. $\to$-sentences can't be: whether an $\to$-conditional gets value 0 or $\frac{1}{2}$ depends on prior values for its antecedent and consequent.)

It is evident that for each world $z$ based on $M$, every $L_0$-sentence gets value 1 in $z$ if it is true in $M$, and 0 in $z$ if it is false in $M$. (Similarly for $L_0^+$-sentences and $M^+$.) So since $\Psi(M)$ will be a product of worlds based in $M$, we have as promised that such ground sentences will always have value $\mathbf{1}_M$ or $\mathbf{0}_M$, depending on whether they are true in $M$.

**§6. The Kripkean construction in this setting.** The task of this section is to generate, for each $\to$-booster $g$ and $\rhd$-valuation $v$, a privileged 3-valued extension $T_{g,v}$ for 'True' which is "the appropriate one given $g$ and $v$". I'll use Kripke's minimal fixed point construction. (It isn't really essential to stick to the *minimal* fixed point, but for convenience I will.) The only novelty, and it's slight, is that the construction will be given in the context provided in the last section, where the values for $\to$-conditionals vary and yet do not reduce to $\supset$. It is only because of special features of the rules for $\to$-conditionals that this will work.

The construction proceeds in stages: for each ordinal $\sigma$, we define a 3-valued extension $T_{g,v,\sigma}$ for 'True' as follows:

- For all $\sigma$, if o isn't (the Gödel number of) an $L$-sentence then $T_{g,v,\sigma}(o) = 0$
- For all $\sigma$, if o is (the Gödel number of) the $L$-sentence $B$ then $T_{g,v,\sigma}(o)$ is $\begin{cases} 1 & \text{if } (\exists\tau < \sigma)(\forall\upsilon \in [\tau,\sigma))(|B|_{g,v,\upsilon} = 1) \\ 0 & \text{if } (\exists\tau < \sigma)(\forall\upsilon \in [\tau,\sigma))(|B|_{g,v,\upsilon} = 0) \\ \frac{1}{2} & \text{otherwise} \end{cases}$

Here, $|B|_{g,v,\upsilon}$ means the value of $B$ relative to $g$ and $v$ and to the choice of $T_{g,v,\upsilon}$ as the 3-extension of 'True'. (Note that if o is an $L^+$-sentence that isn't an $L$-sentence, $T_{g,v,\sigma}(o)$ is always 0.)

So $T_{g,v,0}$ assigns to each (Gödel number of an) $L$-sentence $B$ the value $\frac{1}{2}$. For each later stage, you look at the values of $B$ at prior stages, as determined by the extended Kleene rules of Section 5, and if these values are eventually 1 prior to $\sigma$, then $T_{g,v,\sigma}$ assigns 1 to $B$, and thus $|True(t)|_{g,v,\sigma} = 1$ whenever $t$ denotes (the Gödel number of) $B$. Similarly for 0.

Following Kripke, we show by induction on complexity that

**(*):** for any $L^+$-sentence $A$ and any $g$ and $v$, if $\tau < \sigma$ and $|A|_{g,v;\tau}$ is 1 or 0 then $|A|_{g,v;\sigma} = |A|_{g,v,\tau}$.

Obviously the presence of the $\rhd$ doesn't interfere with this, since $L^+$-sentences of form $B \rhd C$ don't change in value as $\sigma$ increases. What about sentences of form $B \to C$? They clearly never change in value from 1 to anything else as $\sigma$ increases. And the value of $B \to C$ could only change from 0 to anything else if the value of $B$ changed from 1 to something else or the value of $C$ changed from 0 to something else; that is, a violation of (*) when $A$ is of form $B \to C$ would require a violation for one of the less complex sentences $B$ and $C$. So the induction on complexity still goes through.

Once (*) is established, the 1 and 0 clauses for 'True' can be simplified to simple existential quantifications: if o is (the Gödel number of) the $L$-sentence $B$ then

- $$T_{g,\upsilon,\sigma}(o) \text{ is } \begin{cases} 1 & \text{if } (\exists\tau < \sigma)(|B|_{g,\upsilon,\tau} = 1) \\ 0 & \text{if } (\exists\tau < \sigma)(|B|_{g,\upsilon,\tau} = 0) \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

(I didn't write them that way at the start because there would be an apparent danger of assigning to $T_{g,\upsilon,\sigma}(o)$ both 1 and 0.) So if $t$ denotes the $L$-sentence $B$ then

$|True(t)|_{g,\upsilon,\sigma}$is 1 iff $(\exists\tau < \sigma)(|B_{|g,\upsilon;\tau} = 1)$, and 0 iff $(\exists\tau < \sigma)(|B|_{g,\upsilon;\tau} = 0)$.

Moreover, cardinality considerations show that there is an ordinal after which further increase in $\sigma$ makes no difference. (That "closure ordinal" may depend on $g$ and $\upsilon$, as well as on the underlying model $M$.) The desired $T_{g,\upsilon}$ ("the minimal fixed point") is the one for such sufficiently large $\sigma$. For any $L^+$-sentence $B$, we take $|B|_{g,\upsilon}$ to simply be the value generated by this fixed point 3-extension for 'True'.

The fixed point property gives us that for any $g$ and $\upsilon$, if $t$ is an $L^+$- term that denotes the $L$-sentence $B$ then

$|True(t)|_{g,\upsilon} = 1$ iff $|B|_{g,\upsilon} = 1$, and $|True(t)|_{g,\upsilon} = 0$ iff $|B|_{g,\upsilon} = 0$.

Since there are only three possible values, it follows that $|True(t)|_{g,\upsilon} = \frac{1}{2}$ iff $|B|_{g,\upsilon} = \frac{1}{2}$, and hence

**(K):** $|True(t)|_{g,\upsilon} = |B|_{g,\upsilon}$.

(The $t$ needn't be in $L$ as opposed to $L^+$, but typically for any sentence in $L$ there will be closed terms in $L$ that denote it, and **(K)** applies to them. Also in connection with the quantifier rules, **(K)** delivers what we desire: if $A(x)$ is a formula satisfied only by the $L$-sentence $B$, then $|\exists x(A(x) \wedge True(x)|_{g,\upsilon} = |B|_{g,\upsilon}$ in virtue of an $L^+$- term that denotes $B$.)

Is this enough to ensure naive truth as defined in Section 2? No. Naivety requires that if $t$ is an $L^+$-term that denotes $L$-sentence $A$, and $Y$ is an $L^+$-sentence that results from $L^+$-sentence $X$ by replacing an occurrence of $A$ with $True(t)$ or vice versa, then $|Y|_{g,\upsilon} = |X|_{g,\upsilon}$. **(K)** together with the strong Kleene rules ensures this *if the substitution of $A$ with $True(t)$ is not in the scope of any conditional*; but that isn't enough for naivety.

To get the more general intersubstitutivity that naivety requires, we need an assumption about $g$ and $\upsilon$. What we need is that they are **transparent**. For $\upsilon$ to be transparent requires that

Whenever $t$ is a closed $L^+$- term that denotes an $L$-sentence $B$, and $Y$ is a sentence that results from the sentence $X$ by replacing an occurrence of $B$ with $True(t)$ or vice versa, then $\upsilon(Y \rhd Z) = \upsilon(X \rhd Z)$ and $\upsilon(Z \rhd Y) = \upsilon(Z \rhd X)$.

For $g$ to be transparent is analogous. Then using **(K)**, an easy induction yields:

LEMMA 6.1 (Transparency). *If $g$ and $v$ are transparent then for any $L$-sentence $B$ and any $L^+$-term $t$ that denotes it, if $Y$ results from $X$ by replacing an occurrence of $B$ with True(t) or vice versa, then* $|Y|_{g,\upsilon} = |X|_{g,\upsilon}$.

From this it is clear that all we need to do to ensure that our construction yields naive truth is to ensure that the $g$ and $v$ we construct are transparent.

We're going to need one other lemma about the micro-constructions. If $v$ and $w$ are $\triangleright$-valuations, let $\mathbf{v} \leq_{\mathbf{K}} \mathbf{w}$ mean that for any sentences $A$ and $B$, if $v(A \triangleright B) \in \{0, 1\}$ then $w(A \triangleright B) = v(A \triangleright B)$. Define $\leq_K$ for $\rightarrow$-boosters similarly (though the case where $g(A \rightarrow B)$ is 0 doesn't arise).

LEMMA 6.2 (K-Monotonicity). *If v and w are $\triangleright$-valuations such that $v \leq_K w$, then for any $\rightarrow$-booster g and any $L^+$-sentence A, if $|A|_{g,v} \in \{0, 1\}$ then $|A|_{g,w} = |A|_{g,v}$.*

*Proof.* By an obvious induction on the stages in a Kripke construction, with a sub-induction on complexity. More fully, suppose $v \leq_K w$. It suffices to show that for all $\sigma$ and all $A$,

$(I_{\sigma,A})$ If $|A|_{g,v,\sigma} \in \{0, 1\}$ then $|A|_{g,w,\sigma} = |A|_{g,v,\sigma}$.

In the induction, we assume that $(I_{\tau,B})$ holds for every $<\tau,B>$ where $\tau < \sigma$, and also when $\tau=\sigma$ and $B$ is a subformula of $A$ (counting each $B(t)$ as a subformula of $\forall x Bx$). [There's no need to count $B$ and $C$ as subformulas of $B \triangleright C$, though no harm in doing so; but it's essential to count them as subformulas of $B \rightarrow C$.] Given this assumption, we must show that $(I_{\sigma,A})$. This is straightforward. (In the case where $A$ is of form $True(t)$ where $t$ denotes a sentence $B$ that may be of high complexity, we use the inductive assumption that for any $\tau < \sigma$, $(I_{\tau,B})$ holds for all $B$. In the other cases we use that $(I_{\sigma,B})$ holds for $B$ of lower complexity than $A$, together with the supposition that $v \leq_K w$. Details in footnote.)[21] □

**§7. Fibers.** The goal of this section is to construct, for each $\rightarrow$-booster $g$ based on $M$, a set $\mathbf{R}_{M,g}$ of $\triangleright$-valuations (think of them as the fairly "regular" ones), and a certain privileged member $v_g$ of that set. For each fixed $g$ based on $M$, I'll let $\mathbf{F}_{M,g}$ be $\{< g, v >: v \in \mathbf{R}_{M,g}\}$, and call it "the fiber attached to $g$"; $< g, v_g >$ will be called the "base point of the fiber".

To complete the geometric analogy, in the next section I will construct a certain "base space" $\mathbf{J}_M$ of "regular" $\rightarrow$-boosters $g$, with a privileged one @. The set $\mathbf{F}_M$ of "worlds" will then be the set of worlds in fibers attached to the base space: that is, $\bigcup\{\mathbf{R}_{M,g} | g \in \mathbf{J}_M\}$. So the value space $V_M$ for the multi-valued model $\Psi(M)$ will be a set of functions from this to $\{0, \frac{1}{2}, 1\}$. Since I'll be holding $M$ constant throughout Part **2**, it will be convenient

---

[21] If $A$ is either atomic with predicate in the ground language, or of form $True(t)$ where $t$ denotes something other than a sentence, then $(I_{\sigma,A})$ is obvious, since $|A|_{g,v,\sigma}$ is independent of $v$ and $\sigma$. If $t$ denotes an $L$-sentence $B$, then $|True(t)|_{g,v,\sigma} = 1$ iff $(\exists\tau < \sigma)(|B|_{g,v,\tau} = 1)$. By the induction hypothesis, this implies $(\exists\tau < \sigma)(|B|_{g,w,\tau} = 1)$, and hence $|True(t)|_{g,w,\sigma} = 1$. Analogously when $|True(t)|_{g,v,\sigma} = 0$.
If $A$ is a $\triangleright$-conditional, $(I_{\sigma,A})$ comes from the assumption that $v \leq_K w$.
$|B \rightarrow C|_{g,v,\sigma} = 1$ if and only if $g(B \rightarrow C) = 1$ if and only if $|B \rightarrow C|_{g,w,\sigma} = 1$. If $|B \rightarrow C|_{g,v,\sigma} = 0$ then $g(B \rightarrow C) < 1$ and $|B|_{g,v,\sigma} = 1$ and $|C|_{g,v,\sigma} = 0$; so by induction hypothesis $|B|_{g,w,\sigma} = 1$ and $|C|_{g,w,\sigma} = 0$, which with $g(B \rightarrow C) < 1$ yields that $|B \rightarrow C|_{g,w,\sigma} = 0$.
If $|\neg B|_{g,v,\sigma} = 1$, then $|B|_{g,v,\sigma} = 0$; so by induction hypothesis $|B|_{g,w,\sigma} = 0$, so $|\neg B|_{g,w,\sigma} = 1$. Similarly if $|\neg B|_{g,v,\sigma} = 0$.
If $|\forall x Bx|_{g,v,\sigma} = 1$ then for all $t$, $|B(t)|_{g,v,\sigma} = 1$; so by induction hypothesis, for all $t$, $|B(t)|_{g,w,\sigma} = 1$, so $|\forall x Bx|_{g,w,\sigma} = 1$. If $|\forall x Bx|_{g,v,\sigma} = 0$ then for some $t$, $|B(t)|_{g,v,\sigma} = 0$; so by induction hypothesis, for some $t$, $|B(t)|_{g,w,\sigma} = 0$, so $|\forall x Bx|_{g,w,\sigma} = 0$.
Similarly for $B \wedge C$.

to drop the subscript '$M$' for now, but it's important to remember that the fiber space, and the value space formed from it, depend on $M$. Everything here is model-relative; the construction will give no sense to any model-independent notion of a sentence's value.

The constructions of the fibers and of the base space are fixed point constructions, and they are extremely similar. In this section we see how it works for fibers.

Let a **$\rhd$-chain** be a set $Z$ of nonempty sets of transparent $\rhd$-valuations such that if $S_1, S_2 \in Z$ then either $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$.

Let $g$ be any transparent $\rightarrow$-booster; it will be held fixed throughout the construction of the fiber. (There are plenty of transparent $\rightarrow$-boosters: a trivial example is one that assigns $\frac{1}{2}$ to every $\rightarrow$-conditional, and far more interesting ones will be constructed in Section 8.)

Given any such $Z$ and $g$, let $val_g[Z]$ be the $\rhd$-valuation given by

DEFINITION 7.1. $$val_g[Z](A \rhd B) = \begin{cases} 1 & \textit{if } (\exists S \in Z)(\forall w \in S)(\textit{if } |A|_{g,w} = 1 \textit{ then } |B|_{g,w} = 1) \\ 0 & \textit{if } (\exists S \in Z)(\forall w \in S)(|A|_{g,w} = 1 \wedge |B|_{g,w} = 0) \\ \frac{1}{2} & \textit{otherwise} \end{cases}$$

This is a legitimate definition: nothing can be assigned both 1 and 0 since if a chain contains both $S_1$ and $S_2$, $S_1 \cap S_2$ can't be empty since it must be either $S_1$ or $S_2$, which were required to be nonempty.[22] If $Z$ is $\varnothing$, $val_g[Z]$ assigns $\frac{1}{2}$ to every $\rhd$-conditional.

Obviously since $g$ and all members of $\bigcup Z$ are transparent then $val_g[Z]$ is transparent.

If $Z_1$ and $Z_2$ are chains, let $Z_1 \leq Z_2$ mean: $(\forall S_1 \in Z_1)(\exists S_2 \in Z_2)(S_2 \subseteq S_1)$. (Only small sets matter, and the smaller the sets in a chain the bigger the chain. Clearly if $Z_1 \subseteq Z_2$ then $Z_1 \leq Z_2$. $\leq$ is a preorder, that is, reflexive and transitive.) So:

**Observation on Monotonicity for $\rhd$-chains** If $Z_1 \leq Z_2$ then for any $g$, $\mathrm{val}_g[Z_1] \leq_K \mathrm{val}_g[Z_2]$. ($\leq_K$ was defined near the end of the previous section, just prior to the K-Monotonicity Lemma.)

*Proof.* If $val_g[Z_1](A \rhd B)$ is 1 then for some $S_1 \in Z_1$, $(\forall w \in S_1)$(if $A|_{g,w} = 1$ then $|B|_{g,w} = 1$); and so any subset $S_2$ of $S_1$ has this property too; and some such subset is in $Z_2$. So the 1 clause for $Z_1$ guarantees the 1 clause for $Z_2$. The 0 clause is analogous. □

**Truncation Corollary for $\rhd$-chains** If $S$ is a member of the $\rhd$-chain $Z$, let $Z_S$ (the **$S$-truncation** of Z) be $\{S' \in Z : S' \subseteq S\}$. Then $val_g[Z_S] = val_g[Z]$.

*Proof.* By definition, both $Z_S \leq Z$ and $Z \leq Z_S$. [The first is immediate from the fact that $Z_S \subseteq Z$. For the second: Suppose $S' \in Z$. then (since $S \in Z$) either $S' \subseteq S$ or $S \subseteq S'$; in the first case $S'$ is itself in $Z_S$, in the second case it has a subset that is.] But then the Observation gives that $val_g[Z_S]$ and $val_g[Z]$ each extend the other in the $\leq_K$-order, which means that they are the same. □

For each $g$, I inductively define a sequence of sets $R_{g,\mu}$ of transparent $\rhd$-valuations, as follows. (Think of $R_{g,\mu}$ as the set of valuations that are "regular to level $\mu$, relative to $g$".) Let

DEFINITION 7.2. $R_{g,\mu} = \{val_g[Z] : Z \textit{ is a } \rhd\textit{-chain and } (\forall \xi < \mu)(\exists S \in Z)(S \subseteq R_{g,\xi})\}$.

Note that if $\kappa < \mu$ then $R_{g,\mu} \subseteq R_{g,\kappa}$; and that all members of any $R_{g,\mu}$ are transparent.

[22] The definition doesn't really require that $Z$ be a chain, simply that it be a collection of nonempty sets of transparent valuations such that any two members of $Z$ have a common subset in $Z$. This is relevant to giving a more natural semantics in the case of vagueness, but for simplicity I'll confine myself here to chains.

LEMMA 7.3 (Nontriviality of fiber-stages). *For all g and* $\mu$, $R_{g,\mu} \neq \varnothing$; *indeed, for all g and* $\mu$, $R_{g,\mu}$ *has a member* $\upsilon_{g,\mu}$ *such that* $(\forall w \in R_{g,\mu})(\upsilon_{g,\mu} \leq_K w)$.

*Proof.* Suppose that for all $\kappa<\mu$, $R_{g,\kappa} \neq \varnothing$. Then the set $Z_{g,\mu} =_{\mathrm{df}} \{R_{g,\kappa} : \kappa<\mu\}$ is a $\triangleright$-chain that trivially meets the conditions in the definition of $R_{g,\mu}$; so letting $\upsilon_{g,\mu} =_{\mathrm{df}} val_g[Z_{g,\mu}]$, $\upsilon_{g,\mu} \in R_{g,\mu}$.

Moreover, if $Z$ is any $\triangleright$-chain for which $(\forall\xi < \mu)(\exists S \in Z)(S \subseteq R_{g,\xi})$, $Z_{g,\mu} \leq Z$, so by the Observation on Monotonicity for $\triangleright$-chains, $val_g[Z_{g,\mu}] \leq_K val_g[Z]$, that is, $\upsilon_{g,\mu} \leq_K val_g[Z]$. So by definition of $R_{g,\mu}$, $\upsilon_{g,\mu}$ is its minimum member in the $\leq_K$ ordering. □

THEOREM 7.4 (Fixed points in fiber-construction).

*(i)* *For each* $\rightarrow$*-booster g, there is a nonempty set* $\mathbf{R}_g$ *(the set of "g-regular"* $\triangleright$*-valuations) such that for all sufficiently large* $\mu$, $R_{g,\mu} = \mathbf{R}_g$ .

*(ii)* *Letting* $\mathbf{Z}_g$ *be the set of* $\mathbf{R}_g$*-chains, that is, nonempty* $\triangleright$*-chains of subsets of* $\mathbf{R}_g$*, then*

*(FP)* $\quad \mathbf{R}_g = \{val_g[Z] : Z \in \mathbf{Z}_g\}$.

*(iii)* $\mathbf{R}_g$ *has a minimal member* $\upsilon_g$ *in the ordering* $\leq_K$*, and it is just* $val_g[\{\mathbf{R}_g\}]$*, that is, it is* $val_g[Z]$ *for the chain with* $\mathbf{R}_g$ *as its sole member.*

*Proof.*

(i) By a standard fixed point argument,[23] together with the observation that if $\mathbf{R}_g$ were empty then the $R_{g,\mu}$ it is equal to would have to be as well, in violation of the lemma.

(ii) Suppose $\upsilon \in \mathbf{R}_g$ and $\mathbf{R}_g = R_{g,\mu}$. Then $\upsilon \in R_{g,\mu+1}$, so there is a $\triangleright$-chain $Z$ that includes at least one subset $S$ of $R_{g,\mu}$ (i.e., of $\mathbf{R}_g$) such that $\upsilon = val_g[Z]$; but then the truncation $Z_S$ consists entirely of members of $\mathbf{R}_g$, and by the Truncation Corollary, $\upsilon = val_g[Z_S]$. (Conversely, if $\upsilon = val_g[Z]$ and $Z$ is a chain of subsets of $\mathbf{R}_g$, $\upsilon$ clearly satisfies the requirement for being in each $R_{g,\mu}$ and hence in $\mathbf{R}_g$.)

(iii) $\{\mathbf{R}_g\} \leq Z$ for all $Z \in \mathbf{Z}_g$, so the minimality of $\upsilon_g$ follows from the observation on monotonicity. □

For each $\upsilon \in \mathbf{R}_g$, $\upsilon$ is $val_g[Z]$ for at least one $Z \in \mathbf{Z}_g$; pick one and call it $Z_{g,\upsilon}$. Then given (ii) of the Fixed Point Corollary, we have in effect a modal-like neighborhood semantics in each fiber, with the valuations in $\mathbf{R}_g$ as "worlds": for each world $w$, the "$w$-neighborhoods" are the members of $Z_{g,w}$ (or the supersets of such members, depending on how you use the term 'neighborhood'). For $\upsilon_g$ the sole neighborhood is $\mathbf{R}_g$; but typical "worlds" within this neighborhood have much smaller neighborhoods, to which those worlds needn't belong. (In the construction as I've done it, the neighborhoods of each world are nested, but as remarked in note 22, this is not essential to the construction.)

[23] Details: Let $\mathbf{R}_g$ be the intersection of all the $R_{g,\kappa}$. Trivially, for every $\kappa$, $\mathbf{R}_g \subseteq R_{g,\kappa}$; we need that for all sufficiently large $\kappa$ the reverse inclusion holds. If $c$ is the cardinality of the ground model $M$ (and hence of $L^+$) then (since $3^c = 2^c$) there are only $2^c$ $\triangleright$-valuations; so there must be an ordinal $\mu$ with at most $2^c$ predecessors such that for every $\triangleright$-valuation $\upsilon$, if $\upsilon \notin \mathbf{R}_g$, then $(\exists\kappa \leq \mu)(\upsilon \notin R_{g,\kappa})$. Since the sequence of $R_{g,\kappa}$ is a chain, it follows that for every $\upsilon$, if $\upsilon \notin \mathbf{R}_g$, then $\upsilon \notin R_{g,\mu}$, and indeed that for every $\kappa \geq \mu$, $\upsilon \notin R_{g,\kappa}$.

It's worth making explicit the values that result. In general,

$$|A \rhd B|_{g,v} = \begin{cases} 1 & \text{if } (\exists S \in Z_{g,v})(\forall w \in S)(\text{if } |A|_{g,w} = 1 \text{ then } |B|_{g,w} = 1) \\ 0 & \text{if } (\exists S \in Z_{g,v})(\forall w \in S)(|A|_{g,w} = 1 \wedge |B|_{g,w} = 0) \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

Defining $\blacktriangleright$ from $\rhd$ as in Section 3, $|A \blacktriangleright B|_{g,v}$ has the same 0 clause, but is 1 iff $(\exists S \in Z_{g,v})(\forall w \in S)(|A|_{g,w} \leq |B|_{g,w})$.

In particular, since $Z_{g,v_g}$ is $\{\mathbf{R}_g\}$,

$$(*)\ |A \rhd B|_{g,v_g} = \begin{cases} 1 & \text{if } (\forall w \in \mathbf{R}_g)(\text{if } |A|_{g,w} = 1 \text{ then } |B|_{g,w} = 1) \\ 0 & \text{if } (\forall w \in \mathbf{R}_g)(|A|_{g,w} = 1 \wedge |B|_{g,w} = 0) \\ \frac{1}{2} & \text{otherwise;} \end{cases}$$

and

$$(**)\ |A \blacktriangleright B|_{g,v_g} = \begin{cases} 1 & \text{if } (\forall w \in \mathbf{R}_g)(|A|_{g,w} \leq |B|_{g,w}) \\ 0 & \text{if } (\forall w \in \mathbf{R}_g)(|A|_{g,w} = 1 \wedge |B|_{g,w} = 0) \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

Also, since the $w$ in $\mathbf{R}_g$ are precisely the valuations of form $val_g[Z]$ where $Z$ is a chain of members of $\mathbf{R}_g$, we also have

$$|A \rhd B|_{g,v_g} = \begin{cases} 1 & \text{if } (\forall w \in \mathbf{R}_g)(|A \rhd B|_{g,w} = 1) \\ 0 & \text{if } (\forall w \in \mathbf{R}_g)(|A \rhd B|_{g,w} = 0) \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

and similarly for $\blacktriangleright$.

The following result will prove central:

THEOREM 7.5 (Regularity for fibers). *For any $L^+$ sentence A and any g, $|A|_{g,v_g} = 1$ iff $(\forall w \in \mathbf{R}_g)(|A|_{g,w} = 1)$, and $|A|_{g,v_g} = 0$ iff $(\forall w \in \mathbf{R}_g)(|A|_{g,w} = 0)$.*

*Proof.* The right to left of each 'iff' is trivial, and the left to right holds by virtue of the minimality of $v_g$ ((iii) of the Fixed Point Theorem) together with the K-Monotonicity Lemma (Lemma 6.2). □

This means that a great deal about the fiber at $g$ is determined from its **base point** $< g, v_g >$.

From now on I'll use $|A|_g$ as an abbreviation for $|A|_{g,v_g}$.

COROLLARY 7.6 (to Regularity Theorem).

(a) *If $|A|_g = 1$ then $|A \rhd B|_g = |A \blacktriangleright B|_g = |B|_g$*

(b) *If $|A|_g = 0$ or $|B|_g = 1$ then $|A \rhd B|_g = |A \blacktriangleright B|_g = 1$.*

(c) *If $|B|_g = 0$ then $|A \blacktriangleright B|_g = |\neg A|_g$, and $|A \rhd B|_g \geq |\neg A|_g$. (The inequality will be strict when $|A|_g$ is $\frac{1}{2}$ and for all $v$ in $\mathbf{R}_g$ $|A|_{g,v} < 1$.)*

*Proof.* (a) If $|A|_g = 1$ then by the Regularity Theorem, $(\forall w \in R_g)(|A|_{g,w} = 1)$; so from (*) and (**) above, $|A \rhd B|_g$ and $|A \blacktriangleright B|_g$ are each 1 iff $(\forall w \in \mathbf{R}_g)(|B|_{g,w} = 1)$, and 0 iff $(\forall w \in \mathbf{R}_g)(|B|_{g,w} = 0)$. But by Regularity again, that means $|A \rhd B|_g = 1$ iff $|B|_g = 1$, and similarly for 0; which implies that the same holds for $\frac{1}{2}$. Similarly for $\blacktriangleright$.

(b) Immediate from the Regularity Theorem and the 1 clause of (*) and (**).

(c) The part involving $\blacktriangleright$ follows from (a), and the part involving $\rhd$ follows from that. □

By (a), "modus ponens for $\rhd$ holds at all base points $g$": If $|A|_g = |A \rhd B|_g = 1$ then $|B|_g = 1$. Similarly for $\blacktriangleright$.

Also we have

THEOREM 7.7 (Classical collapse for $\rhd$ and $\blacktriangleright$, preliminary statement). *If* $|A \vee \neg A|_g = 1$ *and* $|B \vee \neg B|_g = 1$ *then for all* $v$ *in* $\mathbf{R}_g$, $|A \rhd B|_{g,v} = |A \blacktriangleright B|_{g,v} = |A \supset B|_{g,v}$.

*Proof.* If $|A|_g = 0$, then for all $v$ in $\mathbf{R}_g$, $|A|_{g,v} = 0$, so for all $v$ in $\mathbf{R}_g$, $|A \rhd B|_{g,v} = |A \blacktriangleright B|_{g,v} = |A \supset B|_{g,v} = 1$. Similarly if $|B|_g = 1$. And if $|A|_g = 1$ and $|B|_g = 0$, then for all $v$ in $\mathbf{R}_g$, $|A|_{g,v} = 1$ and $|B|_{g,v} = 0$, so $|A \rhd B|_{g,v} = |A \blacktriangleright B|_{g,v} = |A \supset B|_{g,v} = 0$. □

**§8. The base space.** The construction of the base space has some analogy to the construction of the fibers. But instead of relativizing the construction to a given $\rhd$-valuation $v$ (which would be the analog of what we did with the fibers), we make use of the prior construction of a privileged $v$ for each $g$, viz. the base point $v_g$. We also make modifications reflecting the fact that our $\rightarrowtail$-boosters are only two-valued.

Let an **$\rightarrowtail$-chain** be a set $P$ of nonempty sets of transparent $\rightarrowtail$-boosters such that if $G_1, G_2 \in P$ then either $G_1 \subseteq G_2$ or $G_2 \subseteq G_1$.

Again, let $|A|_g$ abbreviate $|A|_{g,v_g}$. Let boost$[P]$ be the $\rightarrowtail$-booster given by

DEFINITION 8.1. $boost[P](A \rightarrow B) = \begin{cases} 1 & \textit{if } (\exists G \in P)(\forall g \in G)(\textit{if } |A|_g = 1 \textit{ then } |B|_g = 1) \\ \frac{1}{2} & \textit{otherwise} \end{cases}$

Obviously since all members of $\bigcup P$ and all valuations of form $v_g$ are transparent then boost$[P]$ is transparent.

In analogy to the fiber-construction: If $P_1$ and $P_2$ are $\rightarrowtail$-chains, let $P_1 \leq P_2$ mean: $(\forall G_1 \in P_1)(\exists G_2 \in P_2)(G_2 \subseteq G_1)$. (Again, only small sets matter, and the smaller the sets in a chain the bigger the chain. If $P_1 \subseteq P_2$ then $P_1 \leq P_2$.) So:

**Observation on Monotonicity for $\rightarrowtail$-chains** If $P_1 \leq P_2$ then $boost[P_1] \leq_K boost[P_2]$. [This now means only that the relation assigned by $P_1$ is a subset of that assigned by $P_2$.]

*Proof.* If boost$[P_1](A \rightarrow B)$ is 1 then for some $G_1 \in P_1$, $(\forall g \in G_1)$(if $|A|_g = 1$ then $|B|_g = 1$); and so any subset $G_2$ of $G_1$ has this property too; and some such subset is in $P_2$. □

**Truncation Corollary for $\rightarrowtail$-chains** If $G$ is a member of the $\rightarrowtail$-chain $P$, let $P_G$ ("the $G$-truncation of $P$") be $\{G' \in P : G' \subseteq G\}$. Then boost$[P_G]$ = boost$[P]$.

*Proof.* Analogous to the proof for fibers. □

I inductively define a sequence of sets $J_\alpha$ of transparent $\rightarrowtail$-boosters, as follows. (Think of $J_\alpha$ as the set of boosters that are "regular to level $\alpha$".) Let

DEFINITION 8.2. $J_\alpha$ = *{boost[P] : P is a* $\rightarrowtail$*-chain and* $(\forall \gamma < \alpha)(\exists G \in P)(G \subseteq J_\gamma)$*}.*

Clearly if $\beta<\alpha$ then $J_\alpha \subseteq J_\beta$; and all members of any $J_\alpha$ are transparent.

LEMMA 8.3 (Nontriviality of base stages). *For all* $\alpha$, $J_\alpha \neq \varnothing$; *indeed, for all* $\alpha$, $J_\alpha$ *has a member* $g_\alpha$ *such that* $(\forall h \in J_\alpha)(g_\alpha \leq_K h)$.

*Proof.* Suppose that for all $\beta<\alpha$, $J_\beta \neq \varnothing$. Then the set $P_\alpha =_{df} \{J_\beta : \beta<\alpha\}$ is an $\rightarrow$-chain that trivially meets the conditions in the definition of $J_\alpha$; so letting $g_\alpha =_{df}$ boost$[P_\alpha]$, we have $g_\alpha \in J_\alpha$. Moreover, if $P$ is any $\rightarrow$-chain for which $(\forall\gamma < \alpha)(\exists G \in P)(G \subseteq J_\gamma)$, then $P_\alpha \leq P$, so by the Observation on Monotonicity for $\rightarrow$-chains, boost$[P_\alpha] \leq_K$ boost$[P]$, that is, $g_\alpha \leq_K$ boost$[P]$. So since $J_\alpha$ is the set of members of form boost$[P]$ where $(\forall\gamma < \alpha)(\exists G \in P)(G \subseteq J_\gamma)$, $g_\alpha$ is its minimum member in the $\leq_K$ ordering. □

THEOREM 8.4 (Fixed point of base construction).

(i) *There is a nonempty set* $\mathbf{J}$ *(the set of "regular"* $\rightarrow$*- boosters) such that for all sufficiently large* $\alpha$, $J_\alpha = \mathbf{J}$.

(ii) *Letting* $\mathbf{P}$ *be the set of* $\mathbf{J}$*-chains, that is, nonempty* $\rightarrow$*-chains of subsets of* $\mathbf{J}$*, then*
*(FP)* $\mathbf{J} = \{boost[P] : P \in \mathbf{P}\}$.

(iii) $\mathbf{J}$ *has a minimal member @ in the ordering* $\leq_K$*, and it is just boost[*$\{\mathbf{J}\}$*], that is, it is boost[*$P$*] for the chain with* $\mathbf{J}$ *as its sole member.*

*Proof.*

(i) As for the fiber construction.

(ii) As for the fiber construction, but to repeat: Suppose $h \in \mathbf{J}$ and $\mathbf{J} = J_\alpha$. Then $h \in J_{\alpha+1}$, so there is an $\rightarrow$-chain $P$ that includes at least one subset $G$ of $J_\alpha$ (i.e., of $\mathbf{J}$ ) such that $h =$ boost$[P]$; but then the truncation $P_G$ consists entirely of members of $\mathbf{J}$ and by the Truncation Corollary, $h =$ boost$[P_G]$. (Conversely, if $h =$ boost$[P]$ and $P$ is a chain of subsets of $\mathbf{J}$ , $h$ clearly satisfies the requirement for being in each $J_\alpha$ and hence in $\mathbf{J}$.)

(iii) $\{\mathbf{J}\} \leq P$ for all $P \in \mathbf{P}$, so the minimality of @ follows from the observation on monotonicity. □

For each $g \in \mathbf{J}$ , $g$ is boost$[P]$ for some $P \in \mathbf{P}$; pick one and call it $P_g$. As with the fibers, (ii) of the Fixed Point Corollary means that we have in effect a modal-like neighborhood semantics on the base space, with the boosters in $\mathbf{J}$ as "worlds": for each world $g$ the $g$-neighborhoods are the members of $P_g$ (or their supersets). For @ the sole neighborhood is $\mathbf{J}$; but typical "worlds" within this neighborhood have much smaller neighborhoods, to which those worlds needn't belong.

The values that result will of course not be completely analogous to those in the fiber case. In general,

$$g(A, B) = \begin{cases} 1 & \text{if } (\exists G \in P_g)(\forall h \in G)(\text{if } |A|_h = 1 \text{ then } |B|_h = 1) \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

(I'm using $g(A, B)$ as an alternative notation for $g(A \rightarrow B)$, to lessen the possibility of confusion with the value of $A \rightarrow B$ at $g$.) And so

$$|A \rightarrow B|_{g,v} = \begin{cases} 1 & \text{if } (\exists G \in P_g)(\forall h \in G)(\text{if } |A|_h = 1 \text{ then } |B|_h = 1) \\ 0 & \text{if } (\forall G \in P_g)(\exists h \in G)(|A|_h = 1 \wedge |B|_h = 0) \text{ and } |A|_{g,v} = 1 \text{ and } |B|_{g,v} = 0 \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

(Note that in the last two conjuncts of the 0 clause we have $v$, not $v_g$.)

Call $g$ **reflexive** if $g$ is in every member of $P_g$. In the case where $g$ is reflexive and in addition $v = v_g$, the 0 clause simplifies to $|A|_{g,v} = 1 \wedge |B|_{g,v} = 0$. In particular, since not only is @ reflexive but $P_@$ is $\{\mathbf{J}\}$,

$$|A \to B|_@ = \begin{cases} 1 & \text{if } (\forall h \in \mathbf{J})(\text{if } |A|_h = 1 \text{ then } |B|_h = 1) \\ 0 & \text{if } |A|_@ = 1 \text{ and } |B|_@ = 0 \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

**Observation on Modus Ponens:**

**(1):** If $(\forall h \in \mathbf{J})(|A|_h = 1)$ then $|A \to B|_@$ is 1 iff $(\forall h \in \mathbf{J})(|B|_h = 1)$.
**(2):** If $|A|_@ = 1$ and $|A \to B|_@$ is 1 then $|B|_@ = 1$.

(2) says that "modus ponens for $\to$ holds at the **principal base node** $< @, v_@ >$".

A very important difference between this base construction and the fiber construction is that we do not have a "Regularity Theorem" for the Base Space. (For two reasons. First, when $g \leq_K h$, there's no guarantee that even for the same $v$, $|A|_{g,v} \leq_K |A|_{h,v}$: if $g(B \to C) < 1$ and $h(B \to C) = 1$, and $|B|_{g,v} = |B|_{h,v} = 1$ and $|C|_{g,v} = |C|_{h,v} = 0$, then $|B \to C|_{g,v}$ will be 0 and $|B \to C|_{h,v}$ will be 1. Second, we'd need in any case that $|A|_{g,v_g} \leq_K |A|_{h,v_h}$, and there's no way to get from $g \leq_K h$ to $v_g \leq_K v_h$.)

Not only does the proof fail, it is easy to construct counterexamples, based either on embeddings of $\to$ inside $\rhd$ or on embeddings of truth claims that "depend on" $\to$ inside $\rhd$. To begin with (this isn't the counterexample), let $K$ be a standard Curry sentence for $\to$, constructed by the usual methods to be equivalent to $True(\langle K \rangle) \to \bot$, where $\bot$ is $\neg\top$ (i.e., $\neg\forall x(x = x)$). By transparency, $K$ must be equivalent to $K \to \bot$. $g(K \to \bot)$ is 1 at those $g$ for which there's a $G$ in $P_g$ throughout which $K$ has value less than 1, and $\frac{1}{2}$ otherwise; $|K \to \bot|_g$, which is $|K|_g$, must be the same since the assumption that it was 0 would lead to contradiction. Since $@ \in P_@$, $|K|_g$ must have value $\frac{1}{2}$ when $g$ is @, and must have value 1 for some other $g \in \mathbf{J}$.

Now let $K^*$ be the sentence $K \rhd \neg K$. At those $g$ for which $|K|_g = 1$, $|K^*|_{g,v}$ is 0 for any $v$ in $\mathbf{R}_g$; and at those $g$ for which $|K|_g < 1$, $|K^*|_{g,v}$ is 1 for any $v$ in $\mathbf{R}_g$. This is obviously a counterexample to "Base-Regularity": $K^*$ has value 1 at the principal node $< @, v_@ >$, but has value 0 at many other base nodes.

Indeed it is an especially interesting example, in that it's a sentence for which excluded middle holds at every node (so $\top \to (K^* \vee \neg K^*)$ holds at every node), and yet unlike the sentences in the ground model it varies in value from node to node, even base node to base node (so neither $\top \to K^*$ nor $\top \to \neg K^*$ holds at the principal base node). I take it that we should not count such sentences as "behaving classically", so the example illustrates that in this theory obeying excluded middle does not suffice for behaving classically. What does suffice is obeying the law $(\top \to A) \vee (\top \to \neg A)$, which we might call **strong excluded middle**;[24] more on this in the next section.

24 So one way to define the classicality predicate mentioned near the end of Section 4 is as '$True(F(x)) \vee \neg True(F(x))$', where '$F(x)$' denotes the result of prefixing $x$ with '$\top \to$'. So defined, the classicality predicate is highly nonbivalent: while it's value is 1 when applied to formulas for which the claim of strong excluded middle has value 1, it's value is $\frac{1}{2}$ rather than 0 for every other formula. One can get predicates that are closer to bivalent by employing iterations of a determinately operator, which can be defined using '$\to$' in the manner given in Field (2008). (Transfinite iterations are also definable, using 'True' in the limit stages.) The more times one iterates 'determinately' in the definition, the closer to bivalence one gets. But one can never get a fully bivalent classicality predicate in this manner; which is good, since otherwise a sentence that asserts of itself that it is not both classical and true would be paradoxical even within the logic.

## Part 3. VALIDITY AND LAWS

**§9. Validity concepts, conditional introduction theorems, structural rules, reasoning by cases, and classical collapse.** Let $L$ be a language (with the formation rules we've been considering) with an expressively adequate ground fragment $L_0$; and let $M$ be a minimally adequate classical model for its ground fragment. In Part **2** I have constructed a multi-valued model $\Psi(M)$ based on $M$, whose values are functions assigning a member of $\{0, \frac{1}{2}, 1\}$ to each node in $\mathbf{F}_M$ (i.e., in $\{< g, v >: g \in \mathbf{J}$ and $v \in \mathbf{R}_{M,g}\}$). For any assignment function $s$ for $M$, let's say that a formula $A$ of the full language $L$ is **<$M$, $s$>-designated** if in $M$ relative to $s$, its value at the principal base node is 1. Equivalently, if its value is 1 throughout the principal fiber $\{< @, v >: v \in \mathbf{R}_{M,@}\}$.

Call a formula **<$M$, $s$>-super-designated** if its value (relative to $M$ and $s$) is 1 at all nodes (equivalently, all base nodes). While for a wide class of sentences designatedness requires super-designatedness, the example of $K^*$ from the last section shows that this is not invariably so.

For sentences we can simply speak of $M$-designatedness and $M$-superdesignatedness, since the assignment function doesn't matter.

It might seem natural to base a model-theoretic explication of validity on the notion of super-designatedness: e.g. to regard a sentence as valid if and only if for every model $M$ it comes out super-designated, and to regard an inference between sentences as valid if and only if for every model where the premises are super-designated, so is the conclusion. But I think this would have a high cost: reasoning by cases would not come out as an acceptable meta-rule. Indeed we'd have cases where a disjunction was counted "valid" even though each disjunct implied absurdities. We can see this from the example of $K^*$ at the end of the previous section. $K^* \vee \neg K^*$ is super-designated in each model. But in no model is either $K^*$ or its negation super-designated; so the inference from $K^*$ to any absurdity one chooses would come out vacuously valid, and similarly for the inference from $\neg K^*$ to the same absurdity. I think we should be deeply dissatisfied with any account on which a disjunction of two claims comes out logically true and yet each of the disjuncts logically implies that I'm a cheese sandwich.

Designatedness does not have the problem: there are models in which $K^*$ is designated (indeed, every model has this feature), so on the definition of validity (and implication) I'm about to give, $K^*$ does not imply that I'm a cheese sandwich (indeed it doesn't imply anything that isn't itself designated).

DEFINITION 9.1. *If $\Gamma$ is a set of formulas of $L$, and $B$ a formula of $L$, call the inference from $\Gamma$ to $B$* ***model-theoretically valid****, and write $\Gamma \models B$, if for every minimally adequate classical model $M$ of the ground fragment $L_0$ of $L$, and every assignment function $s$ for $M$, if all members of $\Gamma$ are <$M$, $s$>-designated then $B$ is <$M$, $s$>-designated. (I'll take "$\Gamma$* ***model-theoretically implies*** *$B$" to be another way of saying that the inference from $\Gamma$ to $B$ is model-theoretically valid.)*

For present purposes I propose to take model-theoretic validity, so defined, as an explication of validity (for quantificational languages with 'True' and the two conditionals).[25] My primary interest will be in proving that all inferences *of certain forms* are (model-theoretically) valid: those forms are **valid schemas,** or **laws.**

[25] This is not quite my ultimately preferred view. On philosophical grounds I prefer to think of validity as a computationally tractable relation, in the sense that for any recursively enumerable

It will be convenient to introduce a notion I'll call **M-validity** ($\Gamma \vDash_M B$):

$\Gamma \vDash_M B$ if and only if for every assignment function $s$ for $M$, if all members of $\Gamma$ are $<M, s>$-designated then $B$ is $<M, s>$-designated.

This really isn't validity in any interesting sense: any decent model-theoretic account of validity quantifies over a wide range of models, it isn't restricted to a single model. But in proving validity it suffices to prove $M$-validity in a way that manifestly doesn't depend on the underlying ground model $M$, and that's what I'll do. In working with a specific model $M$, we can introduce the extension $L_M^+$ discussed before, and then confine our attention to inferences among $L_M^+$*-sentences*: now taking $\Gamma$ to be a set of $L_M^+$-sentences and $B$ an $L_M^+$-sentence, there is no loss of information in simply saying

$\Gamma \vDash_M B$ if and only if, if all members of $\Gamma$ are $M$-designated then $B$ is $M$-designated.

That is, letting $|\Gamma|_{M,g,v}$ abbreviate $\min\{|A|_{M,g,v} : A \in \Gamma\}$, and $|\Gamma|_{M,g}$ abbreviate $|\Gamma|_{M,g,v_g}$:

**(*):** $\Gamma \vDash_M B$ if and only if (if $|\Gamma|_{M,@} = 1$ then $|B|_{M,@} = 1$);

or equivalently,

$$\Gamma \vDash_M B \text{ if and only if } (\forall v \in \mathbf{R}_@)(\text{if } |\Gamma|_{M,@,v} = 1 \text{ then } |B|_{M,@,v} = 1).$$

This is the notion whose properties I will primarily be investigating. Usually I'll drop the subscript '$M$'. This is potentially dangerous, since it might lead us to forget that we're working with a model-dependent notion rather than full model-theoretic validity; but I think I've stressed the distinction enough to make the risk minimal.

---

set $\Gamma$ of formulas, the set of formulas $B$ that validly follow from it should also be recursively enumerable. Model-theoretical validity as defined above won't satisfy that: indeed, any hope for that was given up from the start, when I restricted to minimally adequate classical models (which, you'll recall, were required to be $\omega$-models).

But even if, like me, you prefer a computationally tractable validity relation, the notion of model-theoretic validity as defined above is highly relevant for the purposes of this paper. For the question of this paper is:

**(Q):** Can we have a logic with naive truth and reasonable laws, including reasonable laws of restricted quantification?

The logic needs to be strong enough so as to include the laws we want (e.g., of naive truth and restricted quantification). It also needs to not be so strong as to declare valid any inferences we strongly want not to be valid. For instance, we certainly don't want the inference from my living in New York to the defeat of the Paris Commune to turn out valid; the conservativeness criterion of Section 4 was imposed precisely so as to guarantee against such an extreme sort of excess strength.

My view is that model-theoretic validity, in the sense defined above, is highly relevant to (Q). What I ultimately want is an axiomatic system, consisting of axioms positing some basic validities together with rules of validity-preservation. As long as

**(R):** The axioms declare valid only inferences that are model-theoretically valid, and the rules are ones that preserve model-theoretic validity,

then any theorem of such a system will declare valid only inferences that are model-theoretically valid, and so the conservativeness criterion will be met.

In short, we are free to choose whatever basic validities and rules of validity-preservation we like, provided that condition **(R)** is met. What is relevant to this paper is only that the inferences we care about are available to be chosen, that is, are model-theoretically valid. How precisely to decide on which of the other model-theoretic validities should count as genuine logical validities (rather than, for example, as mathematical truths) needn't here concern us.

It is easily seen that both modus ponens for $\triangleright$ and modus ponens for $\rightarrow$ are valid ($M$-valid for each $M$), in the sense of ($M$-)validity I've defined. I remarked before that in any naive truth theory meeting extremely minimal conditions, we can't have conditionals that obey both modus ponens and unrestricted conditional proof. Indeed, Curry's paradox shows that modus ponens requires failure of conditional proof *even in the special case when there are no side premises* (or none beyond the arithmetic used to construct self-referential sentences). But there are nonetheless limited versions of conditional proof that hold.

In the case of $\rightarrow$-Introduction, I introduce the notion of universal validity (and universal $M$-validity). If $\Gamma$ is a set of $L_M^+$-sentences and $B$ an $L_M^+$-sentence, define the **universal M-validity** of the inference from $\Gamma$ to $B$ by

**(**):** ($\Gamma \vDash_{M,univ} B$ iff $(\forall g \in \mathbf{J}_M)$(if $|\Gamma|_{M,g} = 1$ then $|B|_{M,g} = 1$)

(where $|\Gamma|_{M,g} = 1$ means that for all $A \in \Gamma$, $|A|_{M,g} = 1$). And if $\Gamma$ is a set of $L$-formulas and $B$ an $L$-formula, call the inference from $\Gamma$ to $B$ **universally valid** ($\Gamma \vDash_{univ} B$) if for every (acceptable) model $M$ and every scheme $s$ for substituting $L_M^+$-names for free variables in $\Gamma$ and $B$, $\Gamma_s \vDash_M B_s$.

Even for 0-premise inferences (i.e., formulas), universal validity is stronger than validity, because of the lack of a regularity theorem on the base space. (In the special case of $\rightarrow$-conditionals, validity implies universal validity.) For $B$ to be universally valid (in a given M) is for $\top \rightarrow B$ to be valid (in that M). While we don't have $B \vDash \top \rightarrow B$ (consider $K^*$ from the previous section), we do have that if $B$ is universally valid so is $\top \rightarrow B$ (and hence that $\top \rightarrow B \vDash \top \rightarrow (\top \rightarrow B)$).

The following meta-theorem (which holds relative to any $M$, as well as generally) is trivial, but worth stating since it will be the standard way of proving the validity of $\rightarrow$-conditionals.

PROPOSITION 9.2 (Limited $\rightarrow$-Introduction).

*(1) If $B \vDash_{univ} C$ then $\vDash_{univ} B \rightarrow C$ (and hence $\vDash B \rightarrow C$)*

*(2) More generally, let $\top \rightarrow \Gamma$ abbreviate $\{\top \rightarrow A : A \in \Gamma\}$. Then for finite $\Gamma$, if $\Gamma, B \vDash_{univ} C$ then $\top \rightarrow \Gamma \vDash_{univ} B \rightarrow C$ (and hence $\top \rightarrow \Gamma \vDash B \rightarrow C$).*

*Proof.* (1) is a special case of (2), so let's just prove the latter. Consider any $g \in J$ for which $|\top \rightarrow \Gamma|_g = 1$. Then for every $A_i$ in $\Gamma$ there is a $G_i$ in $P_g$ such that $(\forall h \in G_i)(|A_i|_h = 1)$; the $G_i$ form a finite subchain of $P_g$, so there is a smallest one, call it $G$. It's in $P_g$, and $(\forall h \in G)(|\Gamma|_h = 1)$. But if $|B \rightarrow C|_g < 1$, then since $G \in P_g$, there must be an $h \in G$ such that $|B|_h = 1$ and $|C|_h < 1$. Since also $|\Gamma|_h = 1$, it can't be the case that $\Gamma, B \vDash_{univ} C$. □

The 'if ... then' in the nonparenthetical parts of (1) and (2) can in fact be replaced by 'if and only if'.[26]

One can *not* generalize (1) to include side formulas, that is, we do *not* have that if $\Gamma, B \vDash_{univ} C$ then $\Gamma \vDash B \rightarrow C$.

**Example:** For any $A$ and $C$, $A, \neg A \vDash_{univ} C$ (since there are no $g$ where $|A|_g$ and $|\neg A|_g$ are both 1); so the purported generalization would require that for all $A$ and $C$,

[26] Suppose that not($\Gamma, B \vDash_{univ} C$); then for some $h \in \mathbf{J}$, $|\Gamma|_h = |B|_h = 1$ and $|C|_h < 1$. $\{\{h\}\}$ is an $\rightarrow$-chain; let $g$ be the $\rightarrow$-booster it generates. Then $|\top \rightarrow \Gamma|_g = 1$ and $|B \rightarrow C|_g < 1$.

$A \vDash \neg A \rightarrow C$. But let $A$ be the sentence $K^*$ from the end of Section 8, and $C$ be $\perp$. $|K^*|_@ = 1$, but there are $g \in \mathbf{J}$ for which $|\neg K^*|_g = 1$, so $|\neg K^* \rightarrow \perp|_@$ isn't 1 (indeed, is $1/2$). □

Might what I'm calling universal validity better be called validity (and what I'm calling validity be called sub-validity)? The Limited $\rightarrow$-Introduction Theorem shows that this would have the advantage of yielding conditional proof in the case where there are no side formulas. Inevitably, this would mean that restrictions would be required on modus ponens. I'll give two examples.

**Example I:** Let $A$ be $K$, the Curry sentence for $\rightarrow$, discussed in the previous section (and used in the definition of $K^*$), and $B$ be $\perp$. As observed earlier, there are base nodes $g$ at which $K$ has value 1; and since it's equivalent to $K \rightarrow \perp$, both $K$ and $K \rightarrow \perp$ have value 1 at such $g$, even though $\perp$ obviously doesn't.

**Example II:** Let $A$ be $K^*$ and $B$ be $\perp$. Let $h$ be any world where $K$ is 1; then $|K^*|_h = 0$. Let $g$ be boost($\{\{h\}\}$). By the valuation rules, $|K \rightarrow \perp|_g$ is 0 and $|K^* \rightarrow \perp|_g = 1$. But the first means that $|K|_g$ is 0 and hence $|K^*|_g$ is also 1.

In both these examples, there are nodes where $A$ and $A \rightarrow B$ have value 1 and $B$ not only has value less than 1, it has value 0. That means that these examples show not only that modus ponens for $\rightarrow$ is not universally valid, but also that the rule

**Contra-MP:** $A, \neg B \vDash \neg(A \rightarrow B)$

is not universally valid. (Both this rule and modus ponens do preserve universal validity, but that is much weaker.[27])

Is the fact that these rules are not universally valid a decisive reason for using the term 'valid' as I have, rather than for universal validity? Perhaps not: it might be thought worthwhile to restrict modus ponens and Contra-MP a bit to get more conditional proof. That would still be congenial to the aims of this paper. For since the target laws from Sections 1 and 3 are all $\rightarrow$-conditionals, their validity implies their universally validity; using universal validity as the explication of 'valid' would still lead to a vindication of the restricted quantifier laws, though in a slightly different overall logic. (Similarly, if anyone is unconvinced that reasoning by cases should be a desideratum and thinks that validity should involve the preservation of super-designatedness, it shouldn't matter: since the target laws from Sections 1 and 3 are all 0-premise inferences, their universal validity means that they preserve super-designatedness. All three notions of validity coincide in the case of the validity of $\rightarrow$-conditionals.)

Another technically useful notion is **super-validity**:

**(***):** $\Gamma \vDash_{M,super} B$ if and only if for every $<g, v>$ in $\mathbf{F}_M$, if $|\Gamma|_{M,g,v} = 1$ then $|B|_{M,g,v} = 1$.

---

27 We don't even have the meta-rule that if $\Gamma \vDash_{univ} B$ and $\Gamma \vDash_{univ} B \rightarrow C$ then $\Gamma \vDash_{univ} C$, when $\Gamma \neq \varnothing$. Taking $\Gamma$ to be $\{B\}$, that would yield the meta-rule that if $B \vDash_{univ} B \rightarrow C$ then $B \vDash_{univ} C$; taking $B$ to be $K$ and $C$ to be $\perp$, we see that this fails. (We do have that if $\vDash_{univ} B \rightarrow C$ then $B \vDash_{univ} C$: that's the converse of (1), proved in the previous footnote.) Similarly, we don't even have the meta-rule that if $\Gamma \vDash_{univ} B$ and $\Gamma \vDash_{univ} \neg C$ then $\Gamma \vDash_{univ} \neg(B \rightarrow C)$, when $\Gamma \neq \varnothing$. Taking $\Gamma$ to be $\{B\}$, that would yield the meta-rule that if $B \vDash_{univ} \neg C$ then $B \vDash_{univ} \neg(B \rightarrow C)$; taking $B$ and $C$ both to be the Liar sentence, we see that this fails.

(Super-validity obviously entails universal validity, and for single sentences, super-validity and universal validity coincide; so valid $\rightarrow$-conditionals are super-valid.) The appropriate introduction theorem here involves the connectives $A \Rightarrow_{nc} B$, defined as $\top \rightarrow (A \rhd B)$, and $A \Rightarrow B$, defined as $\top \rightarrow (A \blacktriangleright B)$. In fact we get biconditionals:

PROPOSITION 9.3 (Limited $\Rightarrow_{nc}$ -Introduction and $\Rightarrow$-Introduction (with converses)). *$B \vDash_{M,super} C$ if and only if $\vDash_M B \Rightarrow_{nc} C$ if and only if $\vDash_{M,univ} B \Rightarrow_{nc} C$.*

*$(B \vDash_{M,super} C$ and $\neg C \vDash_{M,super} \neg B)$ if and only if $\vDash_M B \Rightarrow C$ if and only if $\vDash_{M,univ} B \Rightarrow C$*

*Proof.* The proof of the first biconditional in each line is routine. The second biconditional in each line holds because as noted earlier, $\top \rightarrow B \vDash \top \rightarrow (\top \rightarrow B)$; so in particular (recalling the definition of $\Rightarrow_{nc}$), $B \Rightarrow_{nc} C \vDash \top \rightarrow (B \Rightarrow_{nc} C)$, and so $\vDash_M B \Rightarrow_{nc} C$ implies $\vDash_M \top \rightarrow (B \Rightarrow_{nc} C)$, which is equivalent to $\vDash_{M,univ} B \Rightarrow_{nc} C$. □

Incidentally, not only does $B \Rightarrow_{nc} C$ imply $B \rhd C$ (as is immediate from its definition, given modus ponens for $\rightarrow$), it implies $B \rightarrow C$ as well (indeed universally): for when $|B \rightarrow C|_g < 1$, then $(\forall G \in P_g)(\exists h \in G)(|B|_h = 1 \wedge |C|_h < 1)$, so $(\forall G \in P_g)(\exists h \in G)(|\top|_h = 1 \wedge |B \rhd C|_h < 1)$, so $|B \Rightarrow_{nc} C|_g < 1$.

We also have a notion of **principal-fiber-validity**:

**(****):** $\Gamma \vDash_{M,pf} B$ if and only if $(\forall v \in \mathbf{R}_{M,@})$(if $|\Gamma|_{M,@,v} = 1$ then $|B|_{M,@,v} = 1$).

In disanalogy to the case of universal validity, this coincides with ordinary validity when $\Gamma = \varnothing$, by the Regularity Theorem for fibers. For this notion, we have the following (which also holds relative to any $M$, as well as generally):

PROPOSITION 9.4 (Limited $\rhd$-Introduction and $\blacktriangleright$-Introduction). *If $\Gamma, B \vDash_{pf} C$ then $\Gamma \vDash B \rhd C$.*

*So if $\Gamma, B \vDash_{pf} C$ and $\Gamma, \neg C \vDash_{pf} \neg B$ then $\Gamma \vDash B \blacktriangleright C$.*

*Proof.* (Obviously the second follows from the first.) Suppose that $|B \rhd C|_@ < 1$. Then there is a $v_1 \in \mathbf{R}_@$ for which $|B|_{@,v_1} = 1$ and $|C|_{@,v_1} < 1$. Suppose also that $|\Gamma|_@ = 1$. Then by the Regularity Theorem on fibers, $|\Gamma|_{@,v_1} = 1$. So the inference from $\Gamma$ and $B$ to $C$ isn't pf-valid. □

In another disanalogy to the case of universal validity, one cannot strengthen the conclusions to $\vDash_{pf}$ claims (except when $\Gamma = \varnothing$; but since pf-validity coincides with validity in that case, this is of no interest).

**Example:** Since obviously $A, \top \vDash_{pf} A$, the proposed strengthening would require $A \vDash_{pf} \top \rhd A$. But while the weaker $A \vDash \top \rhd A$ is guaranteed by the above, $A \vDash_{pf} \top \rhd A$ has many counterexamples. For instance, take $A$ to be the "$\rhd$-Curry" $K_\#$ , constructed so as to be equivalent to $True(\langle K_\# \rangle) \rhd \bot$ and therefore to $K_\# \rhd \bot$. For any fiber $g$ including @, $|K_\#|_{g,v_g}$ must be $\frac{1}{2}$. [Reason: since $v_g \in Z_{g,v_g}$, $|K_\# \rhd \bot|_{g,v_g}$ can be 1 only if $|K_\#|_{g,v_g} < 1$, and $|K_\# \rhd \bot|_{g,v_g}$ can be 0 only if $|K_\#|_{g,v_g} = 1$; the only value compatible with $|K_\# \rhd \bot|_{g,v_g} = |K_\#|_{g,v_g}$ is $\frac{1}{2}$.] Since it isn't 1 and is equivalent to $|K_\# \rhd \bot|_{g,v_g}$, there must be a $v$ in $\mathbf{R}_g$ where $|K_\#|_{g,v} = 1$. This requires that for some $S$ in $Z_{g,v}$, $(\forall w \in S)(|K_\#|_{g,w} < 1)$; and that requires that $|\top \rhd K_\#|_{g,v} < 1$. So at such a $v$, $|K_\#|_{g,v} = 1$ and $|\top \rhd K_\#|_{g,v} < 1$. □

I turn now to structural rules, reasoning by cases, and the collapse of both conditionals to $\supset$ in classical contexts.

**Structural Rules:** It is immediate from the definitions that the usual structural rules hold for validity and all its strengthenings (universal validity, pf-validity and super-validity).

That is, for each notion we have both

(1) Weakening: if $\Gamma_1 \vDash B$ and $\Gamma_1 \subseteq \Gamma_2$ then $\Gamma_2 \vDash B$; and
(2) Cut: if $\Gamma, A \vDash B$ and $\Gamma, B \vDash C$ then $\Gamma, A \vDash C$.

And since I've defined $\Gamma \vDash B$ for the case where $\Gamma$ is a *set* (rather than a sequence) of premises, structural permutation and structural contraction are automatic. I will implicitly appeal to all these structural rules in what follows, without special note.

**Reasoning by cases:** For each of these validity notions, we have reasoning by cases even with side formulas, both in the $\vee$-form and the $\exists$-form.

That is,

(1) If $\Gamma, A \vDash C$ and $\Gamma, B \vDash C$ then $\Gamma, A \vee B \vDash C$, and
(2) if $\Gamma, A(x) \vDash C$ when $x$ isn't free in $C$ or any member of $\Gamma$, then $\Gamma, \exists x A(x) \vDash C$.

For instance, the only way we could have a failure of $\Gamma, A \vee B \vDash_{univ} C$ is for there to be a model $M$ and a $g$ in $\mathbf{J}_M$ such that $|\Gamma|_{M,g} = |A \vee B|_{M,g} = 1$ but $|C|_{M,g} < 1$; but then either ($|\Gamma|_{M,g} = |A|_{M,g} = 1$ and $|C|_{M,g} < 1$) or ($|\Gamma|_{M,g} = |B|_{M,g} = 1$ and $|C|_{M,g} < 1$), so at least one of $\Gamma, A \vDash C$ and $\Gamma, B \vDash C$ must fail.

Finally we can state a classical collapse result for all our conditionals:

**Classical collapse:** All conditionals are equivalent to $\supset$ in classical contexts, that is, on the assumption of strong excluded middle for their antecedents and consequents.

That is,

(1) $(\top \rightarrow A) \vee (\top \rightarrow \neg A), (\top \rightarrow B) \vee (\top \rightarrow \neg B) \vDash_{univ} (A \rhd B) \Leftrightarrow (A \supset B)$, and similarly for $\blacktriangleright$; and
(2) $(\top \rightarrow A) \vee (\top \rightarrow \neg A), (\top \rightarrow B) \vee (\top \rightarrow \neg B) \vDash (A \rightarrow B) \Leftrightarrow (A \supset B)$.

*Proof.*

(1) If $|(\top \rightarrow A) \vee (\top \rightarrow \neg A)|_g = 1$ then either $(\exists G \in P_g)(\forall h \in G)(|A|_h = 1)$ or $(\exists G \in P_g)(\forall h \in G)(|A|_h = 0)$. But then by Regularity for fibers, either $(\exists G \in P_g)(\forall h \in G)(\forall w \in \mathbf{R}_h)(|A|_{h,w} = 1)$ or $(\exists G \in P_g)(\forall h \in G)(\forall w \in \mathbf{R}_h)(|A|_{h,w} = 0)$. Similarly for $B$. It clearly follows that $(\exists G \in P_g)(\forall h \in G)(\forall w \in \mathbf{R}_h)(|A \rhd B|_{h,w} = |A \supset B|_{h,w})$, hence $(\exists G \in P_g)(\forall h \in G)(|(A \rhd B) \blacktriangleleft\!\blacktriangleright (A \supset B)|_h)$. Hence $|(A \rhd B) \Leftrightarrow (A \supset B)|_g = 1$.
(2) If $|(\top \rightarrow A) \vee (\top \rightarrow \neg A)|_@ = 1$ then either $(\forall h \in \mathbf{J})(|A|_h = 1)$ or $(\forall h \in \mathbf{J})(|A|_h = 0)$. But then by Regularity for fibers, either $(\forall h \in \mathbf{J})(\forall w \in \mathbf{R}_h)(|A|_{h,w} = 1)$ or $(\forall h \in \mathbf{J})(\forall w \in \mathbf{R}_h)(|A_{|h,w} = 0)$. Similarly for $B$. It clearly follows that $(\forall h \in \mathbf{J})(\forall w \in \mathbf{R}_h)(|A \rightarrow B|_{h,w} = |A \supset B|_{h.w})$, hence $(\forall h \in \mathbf{J})(|(A \rightarrow B) \blacktriangleleft\!\blacktriangleright (A \supset B)|_h)$. Hence $|(A \rightarrow B) \Leftrightarrow (A \supset B)|_@ = 1$. □

Note that since (1) is universal, it generates the valid$\rightarrow$-conditional $[(\top \rightarrow A) \vee (\top \rightarrow \neg A)] \wedge [(\top \rightarrow B) \vee (\top \rightarrow \neg B)] \rightarrow [(A \rhd B) \Leftrightarrow (A \supset B)]$.

Especially in the case of (1), there are related results that derive weaker equivalences from weaker premises (also in universal fashion, so that we get valid $\rightarrow$-conditionals from them). For instance

(1a) $\top \rightarrow [A \vee \neg A \vee B \vee \neg B] \vDash_{univ} (A \blacktriangleright B) \leftrightarrow\!\leftrightarrow (A \supset B)$
(1b) $\top \rightarrow [A \vee \neg A \vee B] \vDash_{univ} (A \rhd B) \leftrightarrow\!\leftrightarrow (A \supset B)$
(1c) $(A \vee \neg A) \wedge (B \vee \neg B) \vDash_{univ} (A \blacktriangleright B) \blacktriangleleft\!\blacktriangleright (A \rhd B) \blacktriangleleft\!\blacktriangleright (A \supset B)$.

*Proof.* (1a) If the premise has value 1 at $g$ then for some $G$ in $P_g$ and all $h$ in $G$, either (i) $|A|_h = 0$ or $|B|_h = 1$, or (ii) $|A|_h = 1$ and $|B|_h < 1$, or (iii) $|B|_h = 0$ and $|A|_h = \frac{1}{2}$. In case (i), regularity gives that either $|A|_{h,w}$ is 0 for all $w$ in $\mathbf{R}_h$ or $|B|_h$ is 1 for all $w$ in $\mathbf{R}_h$, in which case clearly $|A \blacktriangleright B|_h$ and $|A \supset B|_h$ are both 1. In case (ii), regularity gives that $|A|_{h,w}$ is 1 for all $w$ in $\mathbf{R}_h$, and that if $|B|_h$ is 0 then $|B|_{h,w}$ is 0 for all $w$ in $\mathbf{R}_h$; clearly then if $|B|_h$ is 0 then $|A \blacktriangleright B|_h$ and $|A \supset B|_h$ are both 0, and if $\frac{1}{2}$ they are both $\frac{1}{2}$. In case (iii), $|A \blacktriangleright B|_h$ and $|A \supset B|_h$ are both $\frac{1}{2}$. So in all three cases, $|A \blacktriangleright B|_h = |A \supset B|_h$ for all $h$ in $G$; this establishes that the right hand side of (1a) has value 1 at $g$.

(1b) Analogous except without case (iii).

(1c) If the premise has value 1 at $g$ then either (i) $|A|_g = 0$ or (ii) $|B|_g = 1$ or (iii) both $|A|_g = 1$ and $|B|_g = 0$. In each case, regularity gives that they have these values throughout the $g$-fiber. Given this, it's clear that in cases (i) and (ii), $|A \blacktriangleright B|_{g,w}$ and $|A \rhd B|_{g,w}$ and $|A \supset B|_{g,w}$ are all 1 throughout the fiber, and in case (iii) they are all 0 throughout the fiber. So for all $w$ in the $g$-fiber, $|A \blacktriangleright B|_{g,w} = |A \rhd B|_{g,w} = |A \supset B|_{g,w}$; this establishes that the $\blacktriangleleft\!\blacktriangleright$-biconditionals connecting them all have value 1 at $g$. □

**§10. Equivalence and strong implication, and some laws.** Call an $L$-formula $A$ **equivalent** to an $L$-formula $B$, **relative to <$M$, $s$>**, if for every $< g, v > \in \mathbf{F}_M$, $|A|_{M,g,v,s} = |B|_{M,g,v,s}$. Call them equivalent **relative to** $M$ if for every assignment function $s$ for $M$ they are equivalent. Call them **equivalent** (full stop) if they are equivalent for every acceptable ground model $M$.

LEMMA 10.1 (Intersubstitutivity). *Suppose that $X_B$ results from $X_A$ by substituting one or more occurrences of $A$ by $B$. Then for any $M$ and $s$, if $A$ is equivalent to $B$ relative to $M$ and $s$ then $X_A$ is equivalent to $X_B$ relative to $M$ and $s$. (It follows that the same holds with an elimination of the relativization to $s$, or to both $M$ and $s$.)*

*Proof.* By induction on number of substitutions, and for the case of one substitution, by subinduction on the complexity of the embedding of the substituted $A$ within $X_A$. □

COROLLARY 10.2. *If $\vDash_{M,s} A \Leftrightarrow B$, and $X_B$ results from $X_A$ by substituting $B$ for one or more occurrences of $A$, then $\vDash_{M,s} X_A \Leftrightarrow X_B$.*

We can get similar substitution results for other contraposable biconditionals:

(1) For $\blacktriangleleft\!\blacktriangleright$, the analog of the corollary is that if $\vDash_{M,s} A \blacktriangleleft\!\blacktriangleright B$, and $X_B$ results from $X_A$ by substituting $B$ for one or more occurrences of $A$ *that are not in the scope of any* $\rightarrow$, then $\vDash_{M,s} X_A \blacktriangleleft\!\blacktriangleright X_B$; it is established by noting that $\vDash_{M,s} A \blacktriangleleft\!\blacktriangleright B$ corresponds to $A$ and $B$ having the same value on the principal fiber (what we might call **principal-fiber-equivalence**), and proving (in analogy to the proof of the equivalence theorem) that principal-fiber-equivalence is preserved under embedding of all connectives other than $\rightarrow$.

(2) For $\leftrightarrow\!\leftrightarrow$ (defined as $(A \leftrightarrow B) \wedge (\neg A \leftrightarrow \neg B)$), the analog of the corollary is that if $\vDash_{M,s} A \leftrightarrow\!\leftrightarrow B$, and $X_B$ results from $X_A$ by substituting $B$ for one or more

occurrences of $A$ *that are not in the scope of any* $\rhd$, then $\vDash_{M,s} X_A \leftrightarrow\!\leftrightarrow X_B$; it is established by noting that $\vDash_{M,s} A \leftrightarrow\!\leftrightarrow B$ corresponds to $A$ and $B$ having the same value at all base nodes (what we might call **base-equivalence**), and proving (in analogy to the proof of the equivalence theorem) that base-equivalence is preserved under embedding of all connectives other than $\rhd$.

Results for noncontraposable biconditionals, and for one-way conditionals (both contraposable and noncontraposable), are more complicated to state, and I will not take the trouble.

I will say that $A$ **strongly implies** $B$ if and only if $\vDash A \Rightarrow B$ (permitting relativizations to $M$, or to $M$ and $s$, on both sides), and that $A$ strongly implies $B$ **relative to assumption** $C$ if $C \vDash A \Rightarrow B$. Clearly strong implications are universal: if $\vDash A \Rightarrow B$ then $\vDash_{univ} A \Rightarrow B$. More generally, If $\Gamma$ is any set of formulas of form $T \rightarrow C$, and $\Gamma \vDash A \Rightarrow B$, then $\Gamma \vDash_{univ} A \Rightarrow B$.

Here are some very obvious strong equivalences and strong implications (in one case, relative to assumptions):

$\vDash \neg\neg A \Leftrightarrow A$

$\vDash A \wedge B \Leftrightarrow B \wedge A$

$\vDash \neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B)$, and similarly for the other deMorgan laws

$\vDash [A \wedge (B \vee C)] \Leftrightarrow [(A \wedge B) \vee (A \wedge C)]$

$\vDash [A \wedge \forall x Bx] \Leftrightarrow \forall x[A \wedge Bx]$, when$x$isn't free in$A$

$\vDash [A \vee \forall x Bx] \Leftrightarrow \forall x[A \vee Bx]$, when$x$isn't free in$A$

$\vDash \forall x A \Leftrightarrow A$, when$x$isn't free in$A$

$\vDash True(\langle A\rangle) \Leftrightarrow A$

$\vDash A \wedge B \Rightarrow A$

$\top \rightarrow A \vDash (A \wedge B) \Leftrightarrow B$

$\vDash \forall x(Ax \wedge Bx) \Leftrightarrow (\forall x Ax \wedge \forall x Bx)$

$\vDash (\forall x Ax \wedge \exists x Bx) \Rightarrow \exists x(Ax \wedge Bx)$

$\vDash \forall x Ax \Rightarrow At$

$\vDash (A \blacktriangleright B) \Leftrightarrow (\neg B \blacktriangleright \neg A)$

$\vDash [A \blacktriangleright (B \wedge C)] \Rightarrow [(A \blacktriangleright B) \wedge (A \blacktriangleright C)]$(similarly with all $\blacktriangleright$ replaced by $\rhd$ ).

$\vDash [(A \vee B) \blacktriangleright C] \Rightarrow [(A \blacktriangleright C) \wedge (B \blacktriangleright C)]$(similarly with all $\blacktriangleright$ replaced by $\rhd$)

$\vDash [A \rightarrow (B \wedge C)] \Rightarrow [(A \rightarrow B) \wedge (A \rightarrow C)]$(similarly with all $\blacktriangleright$ replaced by $\rhd$)

$\vDash [(A \vee B) \rightarrow C] \Rightarrow [(A \rightarrow C) \wedge (B \rightarrow C)]$(similarly with all $\blacktriangleright$ replaced by $\rhd$)

$\vDash (A \wedge \neg A) \Rightarrow (B \vee \neg B)$.

(Indeed when $Q$ is a noncontingent Liar sentence, $\vDash (A \wedge \neg A) \Rightarrow Q$ and $\vDash Q \Rightarrow (B \vee \neg B)$.)

These are all quite trivial: they can be obtained from obvious super-validities using Limited $\Rightarrow$-Introduction.

Here is an important base-equivalence (which follows from laws below, but is worth stating separately here):

**(W1):** $\vDash (\top \rhd B) \leftrightarrow\!\leftrightarrow B$

It shows a "near-redundancy" of '$\top \rhd$': it is eliminable when not in the scope of a $\blacktriangleright$ or $\rhd$. More generally,

**(W1⁺):** $\models (\top \rightarrow A) \rightarrow ((A \rhd B) \leftrightarrow\!\leftrightarrow B)$.

*Proof.* If $|\top \rightarrow A|_g = 1$ then there is a $Y \in P_g$ such that for all $h \in Y$, $|A|_h = 1$; so by Regularity, for all $< h, v >$ in $\mathbf{F}$ for which $h \in Y$, $|A|_{h,v} = 1$; so for all $h \in Y$,

$$|A \rhd B|_h = \begin{cases} 1 & \text{if for all } v \text{ in } \mathbf{R}_h,\ |B|_{h,v} = 1 \\ 0 & \text{if for all } v \text{ in } \mathbf{R}_h,\ |B|_{h,v}=0 \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

But by regularity that just means that for all $h \in Y$, $|A \rhd B|_h = |B|_h$. So $|(A \rhd B) \leftrightarrow\!\leftrightarrow B|_g = 1$. □

I now list some less trivial laws. I will note which ones are universally valid, since this implies further validities not listed. (In the case of $\rightarrow$-conditional sentences, we've seen that validity, universal validity and super-validity all coincide, but it seemed less distracting to simply mark them as universally valid below.) This is not intended as a systematic exploration of the laws that hold in the system, but is intended to give the general flavor of the impact of the construction, including considerably more than just establishing the laws of restricted quantification given at the start of the paper.

**L1: Modus ponens and related laws for $\rightarrow$**

L1a. $A, A \rightarrow B \models B$

L1b. If $\top \rightarrow \Gamma \models_{univ} B \rightarrow C$ then $\Gamma, B \models_{univ} C$; in particular, if $\models_{univ} B \rightarrow C$ then $B \models_{univ} C$. (So if $\models_{univ} B \rightarrow C$ and $\models_{univ} B$ then $\models_{univ} C$.)

L1c. $\top \rightarrow A, (A \wedge C) \rightarrow B \models_{univ} C \rightarrow B$

*Proof.*

(a) If $|A \rightarrow B|_@ = 1$ then for all $g$ in $\mathbf{J}$, if $|A|_g = 1$ then $|B|_g = 1$; so in particular this holds for $g$=@.

(b) Proved in note 26.

(c) Suppose that $|(A \wedge C) \rightarrow B|_g = 1$ and $|\top \rightarrow A|_g = 1$. Then there is a $Y_1$ in $P_g$ such that for all $h \in Y_1$, if $|A \wedge C|_h = 1$ then $|B|_h = 1$; and a $Y_2$ in $P_g$ such that for all $h \in Y_1$, $|A|_h = 1$. So letting $Y$ be the smaller of them, we have that for all $h \in Y$, if $|C|_h = 1$ then $|B|_h = 1$. So $|C \rightarrow B|_g = 1$. □

**L2: Modus ponens and related laws for $\blacktriangleright$ and $\rhd$**

L2a. $A, A \blacktriangleright B \models_{univ} B$

L2b. $\models_{univ} A \wedge (A \blacktriangleright B) \rightarrow B$

L2c. $A, (A \wedge C) \blacktriangleright B \models_{univ} C \blacktriangleright B$

L2d. $\models_{univ} [A \wedge ((A \wedge C) \blacktriangleright B)] \rightarrow (C \blacktriangleright B)$

All of these hold also with $\rhd$ instead of $\blacktriangleright$.

*Proof.* (I'll prove them for the $\rhd$-forms; the modification for $\blacktriangleright$-forms is obvious.)

(a) For any $g$, if $|A \rhd B|_g = 1$ then for all $w$ in $\mathbf{R}_g$, if $|A|_{g,w} = 1$ then $|B|_{g,w} = 1$; so in particular this holds for $w$=$v_g$, so if $|A \rhd B|_g = 1$ and $|A|_g = 1$ then $|B|_g = 1$.

(b) From (a) by Limited $\rightarrow$-Introduction. (Conversely, (a) follows from (b) by L1b.)

(c) and (d) These are equivalent, by Limited $\rightarrow$-Introduction and L1b. To establish (c): For any $g$, suppose $|A|_g = 1$ and $|(A \wedge C) \rhd B|_g = 1$. By the second equality, for every $v$ in $\mathbf{R}_g$, if $|A \wedge C|_{g,v} = 1$ then $|B|_{g,v} = 1$. By the first equality and Regularity, $|A|_{g,v} = 1$ for all $v$ in $\mathbf{R}_g$. From these we get that for every $v$ in $\mathbf{R}_g$ if $|C|_{g,v} = 1$ then $|B|_{g,v} = 1$, and so $|C \rhd B|_g = 1$. □

Observe that from L1c and L2b we get
$\vDash_{univ} (\top \rhd B) \rightarrow B$,
and indeed the stronger version

**(Mixed Assertion):** $\vDash_{univ} (\top \rightarrow A) \rightarrow [(A \rhd B) \rightarrow B]$;

and similarly with $\blacktriangleright$ for $\rhd$. With L3 below, we get the biconditional versions
$\top \rightarrow A \vDash_{univ} (A \rhd B) \leftrightarrow B$
$\top \rightarrow A \vDash_{univ} (A \blacktriangleright B) \leftrightarrow B$
and their special cases
$\vDash_{univ} (\top \rhd B) \leftrightarrow B$
$\vDash_{univ} (\top \blacktriangleright B) \leftrightarrow B$.

**L3: Connections to $\supset$: positive**

L3a. $\vDash_{univ} (A \supset B) \rightarrow\rightarrow (A \blacktriangleright B)$

L3b. $\top \rightarrow (A \supset B) \vDash_{univ} A \rightarrow\rightarrow B$

*Proof.*

(a) For any $g \in \mathbf{J}$ : If $|A \supset B|_g = 1$ then either $|A|_g = 0$ or $|B|_g = 1$, so by Regularity, either $(\forall v \in \mathbf{R}_g)(|A|_{g,v} = 0)$ or $(\forall v \in \mathbf{R}_g)(|B|_{g,v} = 1)$, so $(\forall v \in \mathbf{R}_g)(|A|_{g,v} \leq |B|_{g,v})$, so $|A \blacktriangleright B|_g = 1$. So $A \supset B \vDash_{univ} A \blacktriangleright B$, and Limited $\rightarrow$-Introduction gives $\vDash_{univ} (A \supset B) \rightarrow (A \blacktriangleright B)$. And if $|A \blacktriangleright B|_g = 0$ then for all $v$ in $\mathbf{R}_g$ and hence in particular at $v_g$ , $|A|_g = 1$ and $|B|_g = 0$, hence $|A \supset B|_g = 0$. So also $\vDash_{univ} \neg(A \blacktriangleright B) \rightarrow \neg(A \supset B)$.

(b) If $|\top \rightarrow (A \supset B)|_g = 1$ then there is a $Y$ in $P_g$ such that for all $h \in Y$, either $|A|_h = 0$ or $|B|_h = 1$, so for every such $h$, $|A|_h \leq |B|_h$, that is, both $\vDash |A \rightarrow B|_g$ and $\vDash |\neg B \rightarrow \neg A|_g = 1$. □

**L4: Connections to $\supset$: negative**

L4a. $\vDash_{univ} \neg(A \supset B) \leftrightarrow \neg(A \blacktriangleright B)$. Similarly for $\rhd$.

L4b. $\vDash_{univ} \neg(A \rightarrow B) \Rightarrow_{nc} \neg(A \supset B)$

L4c. $\neg(A \supset B) \vDash \neg(A \rightarrow B)$

Hence $\neg(A \rhd B) \vDash \neg(A \rightarrow B)$, by L4a.
$\neg(A \supset B)$ is of course equivalent to $A \wedge \neg B$, and it might be more natural to rewrite these using $A \wedge \neg B$.

*Proof.*

(a) $|\neg(A \supset B)|_g = 1$ iff $|A|_g = 1$ and $|B|_g = 0$, iff for all $v$ in $\mathbf{R}_g$, $|A|_{g,v} = 1$ and $|B|_{g,v} = 0$, iff $|A \blacktriangleright B|g = 0$. Limited $\rightarrow$ Introd gives the result.

(b) If $|\neg(A \rightarrow B)|_{g,v} = 1$ then $|A|_{g,v} = 1$ and $|B|_{g,v} = 0$, that is, $|\neg(A \supset B)|_{g,v} = 1$. So the inference is super-valid, so by Limited $\Rightarrow_{nc}$ -Introduction, $\vDash_{univ} \neg(A \rightarrow B) \Rightarrow_{nc} \neg(A \supset B)$.

(c) If $|\neg(A \supset B)|_{@} = 1$, then since @ is in the only member of $P_{@}$, $@(A, B) < 1$. Also $|A|_{@} = 1$ and $|B|_{@} = 0$. These things together give that $|A \rightarrow B|_{@} = 0$. □

As already remarked, the implication (c) isn't universal: it depends on the reflexivity of @.

**L5: Conjunction laws.** Let $\gg$ be any of the conditionals $\rightarrow$, $\rhd$, and $\blacktriangleright$ (or, for that matter $\Rightarrow$ or $\Rightarrow_{nc}$ , or any of the biconditionals).

(L5a) $\vDash_{univ}[(A \gg B) \wedge (A \gg C)] \Rightarrow_{nc} [A \gg (B \wedge C)]$

(L5b) $\vDash_{univ}[(A \blacktriangleright B) \wedge \neg(A \blacktriangleright (B \wedge C))] \rightarrow \neg[A \blacktriangleright C]$, and the same with all $\blacktriangleright$ changed to $\rhd$

(L5c) $\vDash_{univ}[(A \blacktriangleright B) \wedge (A \blacktriangleright C)] \rightarrow\rightarrow [A \blacktriangleright (B \wedge C)]$, and the same with all $\blacktriangleright$ changed to $\rhd$

(L5d) $(A \rightarrow B) \wedge \neg(A \rightarrow (B \wedge C)) \vDash \neg[A \rightarrow C]$

(L5e) $\vDash [(A \rightarrow B) \wedge (A \rightarrow C)] \blacktriangleright [A \rightarrow (B \wedge C)]$

*Proof.*

(a) Given Limited $\Rightarrow_{nc}$ -Introd, we need only that for all $g$ in $\mathbf{J}$ , and all $v$ in $\mathbf{R}_g$, if $|(A \gg B) \wedge (A \gg C)|_{g,v} = 1$ then $|A \gg (B \wedge C)|_{g,v} = 1$.
When $\gg$ is $\blacktriangleright$ : If (for a given $g$ and $v$) $|(A \blacktriangleright B) \wedge (A \blacktriangleright C)|_{g,v} = 1$ then there are $S_1$ and $S_2$ in $Z_{g,v}$ such that for all $w$ in $S_1$, $|A|_{g,w} \leq |B|_{g,w}$, and for all $w$ in $S_2$, $|A|_{g,w} \leq |C|_{g,w}$. Letting $S$ be whichever is smaller, for all $w$ in $S$ both $|A|_{g,w} \leq |B|_{g,w}$ and $|A|_{g,w} \leq |C|_{g,w}$, and so $|A|_{g,w} \leq |B \wedge C|_{g,w}$. So $|A \blacktriangleright (B \wedge C)|_{g,v} = 1$.
Analogously for $\rhd$ , and pretty much analogously for $\rightarrow$ too: If (for a given $g$ and $v$) $|(A \rightarrow B) \wedge (A \rightarrow C)|_{g,v}$ is 1 (which actually is independent of $v$), there are $Y_1$ and $Y_2$ in $P_g$ such that for all $h$ in $Y_1$ if $|A|_h = 1$ then $|B|_h = 1$ and for all $h$ in $Y_2$ if $|A|_h = 1$ then $|C|_h = 1$. So going to the smaller, there's a $Y$ such that for all $h$ in $Y$, if $|A|_h = 1$ then $|B \wedge C|_h = 1$, so $|A \rightarrow (B \wedge C)|_g = 1$, and this value is unchanged for other $v$ in $\mathbf{R}_g$.

(b) Given Limited $\rightarrow$-Introduction, we need only that for all $g$ in $\mathbf{J}$ , if $|A \blacktriangleright B|_g = 1$ and $|A \blacktriangleright (B \wedge C)|_g = 0$ then $|A \blacktriangleright C|_g = 0$. Suppose $|A \blacktriangleright B|_g = 1$. Then for all $w$ in $\mathbf{R}_g$, if $|A|_{g,w} = 1$ then $|B|_{g,w} = 1$. If $|A \blacktriangleright (B \wedge C)|_g = 0$ then $|A|_g = 1$, and either $|B|_g = 0$ or $|C|_g = 0$. But by the previous, $|B|_g$ (i.e., $|B|_{g,v_g}$) can't be 0, since $v_g$ is in $\mathbf{R}_g$. So $|A|_g$ is 1 and $|C|_g$ is 0, and Regularity gives that for all $w$ in $\mathbf{R}_g$, $|A|_{g,w} = 1$ and $|C|_{g,w} = 0$, and so $|A \blacktriangleright C|_g = 0$.

(c) I showed earlier that $\Rightarrow_{nc}$ universally implies $\rightarrow$; so the $\rightarrow$ claim follows from (a), and we need only show that for all $g$ in $\mathbf{J}$ , if $|A \blacktriangleright (B \wedge C)|_g = 0$, then for all $w$ in $\mathbf{R}_g$, $|A|_{g,w} = 1$ and $|B \wedge C|_{g,w} = 0$. In particular the second holds at $v_g$, so either $|B|_{g,v_g} = 0$ or $|C|_{g,v_g} = 0$, so by Regularity either for all $w$ in $\mathbf{R}_g$ $|B|_{g,w} = 0$ or for all $w$ in $\mathbf{R}_g$ $|C|_{g,w} = 0$. Since in addition for all $w$ in $\mathbf{R}_g$ $|A|_{g,w} = 1$, we have that either $|A \blacktriangleright B|_g = 0$ or $|A \blacktriangleright C|_g = 0$.

(d) Suppose $|A \rightarrow B|_{@} = 1$. Then for all $h$ in $\mathbf{J}$ , if $|A|_h = 1$ then $|B|_h = 1$. If $|A \rightarrow (B \wedge C)|_{@} = 0$ then $|A|_{@} = 1$, either $|B|_{@} = 0$ or $|C|_{@} = 0$, and by the previous it must be that $|A|_{@} = |B|_{@} = 1$ and $|C|_{@} = 0$. Given the reflexivity of @, this is enough to establish that $|A \rightarrow C|_{@} = 0$.

(e) The corresponding $\rhd$ claim follows from (a), so what's needed in addition is that for all $v$ in $\mathbf{R}_{@}$, if $|A \rightarrow (B \wedge C)|_{@,v} = 0$ then $|(A \rightarrow B) \wedge (A \rightarrow C)|_{@,v} = 0$.

If $|A \to (B \wedge C)|_{@,v} = 0$ then $|A|_{@,v} = 1$, and also $|B \wedge C|_{@,v} = 0$. So either $|A|_{@,v} = 1$ and $|B|_{@,v} = 0$, or $|A|_{@,v} = 1$ and $|C|_{@,v} = 0$. Without loss of generality assume the former. @ is in $\mathbf{J}$, so we have in addition that $@(A, B) < 1$; so $|A \to B|_{@,v} = 0$. □

Note that the validities in (d) and (e) aren't universal: they depend on the reflexivity of @.

**L5*: Disjunction laws.** The disjunctive analogs of L5(a)–(e).

For example, the analog of (a) is that for each conditional $\gg$, $\vDash_{univ} [(A \gg C) \wedge (B \gg C)] \Rightarrow_{nc} [(A \vee B) \gg C]$. The proofs of this and the others are analogous to those of L5.

**L6: Positive conjunctive syllogisms.** Let $\gg$ be as in L5.

$$\text{(L6)} \vDash_{univ} [(A \gg B) \wedge (B \gg C)] \Rightarrow_{nc} (A \gg C)$$

*Proof.* ($\blacktriangleright$ case) We need that for all $g$ and $v$, if $|A \blacktriangleright B|_{g,v} = |B \blacktriangleright C|_{g,v} = 1$ then $|A \blacktriangleright C|_{g,v} = 1$. That's clear: the premises imply that for some $S_1, S_2 \in Z_{g,v}$, $(\forall w \in S_1)(|A|_{g,w} \leq |B|_{g,w})$ and $(\forall w \in S_2)(|B|_{g,w} \leq |C|_{g,w})$; so letting $S$ be the smaller of the two, we have $(\forall w \in S)(|A|_{g,w} \leq |C|_{g,w})$, and so $|A \blacktriangleright C|_{g,v} = 1$.

($\triangleright$ case): Analogous.

($\to$ case): Also closely analogous. We need that for all $g$ and $v$, if $|A \to B|_{g,v} = |B \to C|_{g,v} = 1$ then $|A \to C|_{g,v} = 1$, which since only value 1 is concerned just requires that for all $g$, if $g(A, B) = g(B, C) = 1$ then $g(A, C) = 1$. The premises imply that for some $Y_1, Y_2 \in P_g$, $(\forall h \in Y_1)$(if $|A|_h = 1$ then $|B|_h = 1$) and $(\forall h \in Y_2)$(if $|B|_h = 1$ then $|C|_h = 1$); so letting $Y$ be the smaller of the two, we have $(\forall h \in Y)$(if $|A|_h = 1$ then $|C|_h = 1$), and so $g(A \to C) = 1$. □

An expanded version will prove useful:

**L6⁺:** $\vDash_{univ} [(A \gg B) \wedge (B \wedge D \gg C)] \Rightarrow_{nc} (A \wedge D \gg C)$.

The proof is essentially the same.

**L7: Negative conjunctive syllogisms.**

(L7a) $\vDash_{univ} [(A \blacktriangleright B) \wedge \neg(A \blacktriangleright C)] \Rightarrow_{nc} \neg(B \blacktriangleright C)$, and similarly for $\triangleright$.

(L7b-i) $(A \to B) \wedge \neg(A \to C) \vDash \neg(B \to C)$

(L7b-ii) $\top \to (A \to B) \vDash \neg(A \to C) \to \neg(B \to C)$

*Proof.*

(a) We need that for all $g$ and $v$, if $|A \blacktriangleright B|_{g,v} = 1$ and $|A \blacktriangleright C|_{g,v} = 0$ then $|B \blacktriangleright C|_{g,v} = 0$. The premises imply that for some $S_1, S_2 \in Z_{g,v}$, $(\forall w \in S_1)(|A|_{g,w} \leq |B|_{g,w})$ and $(\forall w \in S_2)(|A|_{g,w} = 1$ and $|C|_{g,w} = 0)$; so letting $S$ be the smaller of the two, we have $(\forall w \in S)(|B|_{g,w} = 1$ and $|C|_{g,w} = 0)$, and so $|B \blacktriangleright C|_{g,v} = 0$.

(b-i) The premise requires that $@(A, B) = 1$ and $@(A, C) < 1$ and $|A|_{@,v} = 1$ and $|C|_{@,v} = 0$. From L6 we know that first two require $@(B, C) < 1$. And $@(A, B) = 1$ and $|A|_{@,v} = 1$ imply $|B|_{@,v} = 1$, by the reflexivity of @; so with $|C|_{@,v} = 0$ we get $|B \to C|_{@} = 0$.

(b-ii) The premise requires that for all $g \in \mathbf{J}$, $g(A, B) = 1$; which requires that for all $h$ in $\mathbf{J}$, if $|A|_h = 1$ then $|B|_h = 1$. So for any $G \subseteq \mathbf{J}$, if $(\forall h \in G)$

(if $|B|_h = 1$ then $|C|_h = 1$) then $(\forall h \in G)$(if $|A|_h = 1$ then $|C|_h = 1$); so for any $g \in \mathbf{J}$, if $g(B, C) = 1$ then $g(A, C) = 1$. Now for any $g \in \mathbf{J}$ , if $|A \rightarrow C|_g = 0$ then $g(A, C) < 1$, $|A|_g = 1$, and $|C|_g = 0$; so by what's been shown we have $g(B, C) < 1$, $|B|_g = 1$, and $|C|_g = 0$; so $|B \rightarrow C|_g = 0$. □

Note that the implications (b-i) and (b-ii) aren't universal: they depend on the reflexivity of @.

**L8: Exportation.**

(L8a) $\models_{univ} [A \wedge B \Rightarrow_{nc} C] \rightarrow [A \rightarrow (B \rhd C)]$

(L8b) $\models_{univ} [A \wedge B \Rightarrow C] \rightarrow [A \rightarrow (B \blacktriangleright C)]$.

*Proof.* For (a), we need that for any $g \in \mathbf{J}$ , if $|A \wedge B \Rightarrow_{nc} C|_g = 1$ then $|A \rightarrow (B \rhd C)|_g = 1$. Suppose the former. Then for some $G$ in $P_g$, $(\forall h \in G)(|A \wedge B \rhd C|_h = 1)$. But by (the $\rhd$ version of) L2c, this implies that $(\forall h \in G)$(if $|A|_h = 1$ then $|B \rhd C|_h = 1$), which means that $|A \rightarrow (B \rhd C)|_g = 1$. The (b) case is analogous. □

**L9: Suffixing and Prefixing.** Let $\gg$ be as in L5.

(L9a) $\models_{univ} (A \gg B) \rightarrow [(B \gg C) \rhd (A \gg C)]$ and $\models_{univ} (A \gg B) \rightarrow [(C \gg A) \rhd (C \gg B)]$.

(L9a$^+$) Indeed when $\gg$ is $\rhd$ or $\blacktriangleright$, we have $\models_{univ} (A \gg B) \rightarrow [(B \gg C) \blacktriangleright (A \gg C)]$, and when it's $\blacktriangleright$ we also get $\models_{univ} (A \gg B) \rightarrow [(C \gg A) \blacktriangleright (C \gg B)]$.

(L9b) $\models_{univ} (A \Rightarrow B) \rightarrow [(B \blacktriangleright C) \rightarrow\!\!\rightarrow (A \blacktriangleright C)]$ and $\models_{univ} (A \Rightarrow B) \rightarrow [(C \blacktriangleright A) \rightarrow\!\!\rightarrow (C \blacktriangleright B)]$.
Similarly, $\models_{univ} (A \Rightarrow_{nc} B) \rightarrow [(B \rhd C) \rightarrow\!\!\rightarrow (A \rhd C)]$ and $\models_{univ} (A \Rightarrow_{nc} B) \rightarrow [(C \rhd A) \rightarrow (C \rhd B)]$.

(L9c) $\models_{univ} [\top \rightarrow (A \rightarrow B)] \rightarrow [(B \rightarrow C) \rightarrow (A \rightarrow C)]$ and $\models_{univ} [\top \rightarrow (A \rightarrow B)] \rightarrow [(C \rightarrow A) \rightarrow (C \rightarrow B)]$

Regarding (a$^+$), (b) and (c), note that we wouldn't expect the $\blacktriangleright$ version of prefixing for a noncontraposable conditional: if $Q$ is the Liar sentence, $Q \gg \bot$ gets value 1 everywhere when $\gg$ doesn't contrapose; but $(\top \gg Q) \blacktriangleright (\top \gg \bot)$ requires $\neg(\top \gg \bot) \rhd \neg(\top \gg Q)$, which will have value $\frac{1}{2}$ at every node since at every node $\neg(\top \gg \bot)$ has value 1 and $\neg(\top \gg Q)$ has value $\frac{1}{2}$.

*Proof.*

(a) Applying Exportation L8a (or the rule that results from applying modus ponens to it) to the version of Positive Conjunctive Syllogism for the appropriate conditional, we get the Suffixing claim $(A \gg B) \rightarrow [(B \gg C) \rhd (A \gg C)]$. If we reverse the conjuncts in Positive Conjunctive Syllogism before applying Exportation, we get $(B \gg C) \rightarrow [(A \gg B) \rhd (A \gg C)]$, which upon relettering gives the desired Prefixing result.

(a$^+$) Similarly, when $\gg$ is $\rhd$ or $\blacktriangleright$, applying Exportation to Negative Conjunctive Syllogism in form L7a, we get $(A \gg B) \rightarrow [\neg(A \gg C) \rhd \neg(B \gg C)]$, which with L5a and L9a gives $(A \gg B) \rightarrow [(B \gg C) \blacktriangleright (A \gg C)]$. A special case of that result is $(\neg B \gg \neg A) \rightarrow [(\neg A \gg \neg C) \rhd (\neg B \gg \neg C)]$, which when $\gg$ is the contraposable $\blacktriangleright$ yields $(A \gg B) \rightarrow [(C \gg A) \blacktriangleright (C \gg B)]$.

(b) and (c): For any $g \in \mathbf{J}$, suppose $|\top \to (A \gg B)|_g = 1$. (Recall that the definitions of $\Rightarrow$ and $\Rightarrow_{nc}$ have this form.) Then for some $G \in P_g$, $(\forall h \in G)(|A \gg B|_h = 1)$. Then

(i) By Positive Conjunctive syllogism, we have that $(\forall h \in G)$(if $|B \gg C|_h = 1$ then $|A \gg C|_h = 1$), and so $|(B \gg C) \to (A \gg C)|_g = 1$; and that $(\forall h \in G)$(if $|C \gg A|_h = 1$ then $|C \gg B|_h = 1$), and so $|(C \gg A) \to (C \gg B)|_g = 1$.

(ii) When $\gg$ is $\rhd$ or $\blacktriangleright$, Negative Conjunctive Syllogism in form L7a yields that $(\forall h \in G)$(if $|A \gg C|_h = 0$ then $|B \gg C|_h = 0$); so $|\neg(A \gg C) \to \neg(B \gg C)|_g = 1$. (i) establishes L9c and much of L9b; what's left is the contrapositive direction of the three $\to\to$ claims. (ii) establishes the contrapositive $\to$ in Suffixing when $\gg$ is $\rhd$ or $\blacktriangleright$. And when $\gg$ is the contraposable $\blacktriangleright$, we get the Prefixing claim from the Suffixing claim. □

An argument analogous to that in L9a also yields the related laws
$\vDash_{univ} \neg(A \gg C) \to ((A \gg B) \rhd \neg(B \gg C))$.[28]

**L10: Strong Explosion.**

$\vDash_{univ} A \wedge \neg A \Rightarrow_{nc} B$.

*Proof.* Obvious. □

(Any analog with a contraposable conditional $\gg$ in place of $\Rightarrow_{nc}$ obviously fails: it would lead to $\neg B \gg (A \vee \neg A)$, by deMorgan's and double negation, and hence to excluded middle by modus ponens for $\gg$ when $B$ is $\bot$.)

**L11: Laws of Quantified Conditionals.**

(L11a) $\vDash_{univ} \forall x(Ax \blacktriangleright Bx) \to\to (\forall x Ax \blacktriangleright \forall x Bx)$ and $\vDash_{univ} \forall x(Ax \blacktriangleright Bx) \to\to (\exists x Ax \blacktriangleright \exists x Bx)$; and analogously for $\rhd$.

(L11b) $\forall x(Ax \to Bx) \vDash \forall x Ax \to \forall x Bx$, and $\forall x(Ax \to Bx) \vDash \exists x Ax \to \exists x Bx$

(L11c) $\neg(\forall x Ax \to \forall x Bx) \vDash \neg\forall x(Ax \to Bx)$, and $\neg(\exists x Ax \to \exists x Bx) \vDash \neg\forall x (Ax \to Bx)$

*Proof.* (a) (For $\rhd$:) We need that for any $g$, $|\forall x(Ax \rhd Bx)|_g \leq |\forall x Ax \rhd \forall x Bx|_g$ and $|\forall x(Ax \rhd Bx)|_g \leq |\exists x Ax \rhd \exists x Bx|_g$.

(1) Suppose $|\forall x(Ax \rhd Bx)|_g$ is 1. Then for each $t$ (in $L_M^+$) and each $w$ in $\mathbf{R}_g$, if $|A(t)|_{g,w}$ is 1 then $|B(t)|_{g,w}$ is 1. It follows that for each $w$ in $\mathbf{R}_g$:

(i) if for each $t$ (in $L_M^+$) $|A(t)|_{g,w}$ is 1 then for each $t$ $|B(t)|_{g,w}$ is 1

(ii) and if for some $t$ (in $L_M^+$) $|A(t)|_{g,w}$ is 1 then for some $t$ $|B(t)|_{g,w}$ is 1. So $|\forall x Ax \rhd \forall x Bx|_g$ and $|\exists x Ax \rhd \exists x Bx|_g$ are 1.

(2) Suppose $|\forall x Ax \rhd \forall x Bx|_g = 0$. Then for all $w$ in $\mathbf{R}_g$, $|\forall x Ax|_{g,w} = 1$ and $|\forall x Bx|_{g,w} = 0$; in particular this holds at $v_g$. So there is a $t$ (in $L_M^+$) such that $|A(t)|_g = 1$ and $|B(t)|_g = 0$, and hence (by Regularity) such that at all $w$ in $\mathbf{R}_g$ $|A(t)|_{g,w} = 1$ and $|B(t)|_{g,w} = 0$. So $|A(t) \rhd B(t)|_g = 0$ and so $|\forall x(Ax \rhd Bx)|_g = 0$. The argument for when $|\exists x Ax \rhd \exists x Bx|_g = 0$ is analogous.

[28] To get analogs of this with $\to$ in place of $\rhd$, one needs to beef up the antecedent $\neg(A \gg C)$ to $\top \to \neg(A \gg C)$; moreover, when $\gg$ is $\to$, one must weaken to (nonuniversal) rule form.

(b) Suppose that for every $t$ (in $L_M^+$), $|A(t) \rightarrow B(t)|_@ = 1$. Then for each $t$, and each $g$ in **J**, if $|A(t)|_g = 1$ then $|B(t)|_g = 1$. Reversing quantifiers and distributing the "for each $t$" quantifier, we have that for each $g$ in **J** : if for each $t$ $|A(t)|_g = 1$ then for each $t$ $|B(t)|_g = 1$, and if for some $t$ $|A(t)|_g = 1$ then for some $t$ $|B(t)|_g = 1$. So if $|\forall x Ax|_g = 1$ then $|\forall x Bx|_g = 1$, and if $|\exists x Ax|_g = 1$ then $|\exists x Bx|_g = 1$; so $|\forall x Ax \rightarrow \forall x Bx|_@ = |\exists x Ax \rightarrow \exists x Bx|_@ = 1$.
(c) If $|\forall x Ax \rightarrow \forall x Bx|_@ = 0$ then $|\forall x Ax|_@ = 1$ and $|\forall x Bx|_@ = 0$. So for some $t$, $|At|_@ = 1$ and $|Bt|_@ = 0$. Using reflexivity of @ we have that $|At \rightarrow Bt|_@ = 0$, and result follows. Similarly for second. □

There are corollaries for other defined conditionals, such as $\rightarrow\rightarrow$. For instance

**Observation:** $\forall x(Ax \rightarrow\rightarrow Bx) \vDash \forall x Ax \rightarrow\rightarrow \forall x Bx$.

*Proof.* The premise says that $\forall x(Ax \rightarrow Bx) \wedge \forall x(\neg Bx \rightarrow \neg Ax)$. Applying both parts of L11b we get $(\forall x Ax \rightarrow \forall x Bx) \wedge (\exists x \neg Bx \rightarrow \exists x \neg Ax)$, which is equivalent to $(\forall x Ax \rightarrow \forall x Bx) \wedge (\neg \forall x Bx \rightarrow \neg \forall x Ax)$, that is, to $\forall x Ax \rightarrow\rightarrow \forall x Bx$. □

## Part 4. CONCLUSION

**§11. Restricted quantification again.** It should be fairly obvious that the previous two sections are more than enough to verify my promises in Section 1, including but not limited to the validation of the target laws of restricted quantification; and indeed that it validates their strengthened versions given in Section 3. But let's make it official.

First, I required the standard structural rules, and reasoning by cases ($\vee$-Elimination and $\exists$-elimination); and I required modus ponens for $\rightarrow$. I've established these (and modus ponens for $\triangleright$ and hence $\blacktriangleright$ as well).

Second, I required that the conditionals "collapse to $\supset$ in classical contexts". More precisely, any such conditional is fully equivalent to the corresponding $\supset$-conditional *on the assumption of strong excluded middle for its antecedent and consequent*; that is,

$(\top \rightarrow A) \vee (\top \rightarrow \neg A), (\top \rightarrow B) \vee (\top \rightarrow \neg B) \vDash (A \triangleright B) \Leftrightarrow (A \supset B)$,

and similarly with $\rightarrow$ for $\triangleright$. I've established these. I also required $A \wedge \neg B \vDash \neg(A \rightarrow B)$, which was established in L4c.

Third, I promised a number of laws of universal restricted quantification, when $\forall x(Ax/Bx)$ is defined as $\forall x(Ax \blacktriangleright Bx)$. The promised laws are all $\rightarrow$-conditionals (or $\rightarrow\rightarrow$-conditionals), so as noted before, demonstrating their validity is in effect verifying their super-validity. The laws, stated in terms of $\blacktriangleright$, were as follows:

**(I):** $\vDash \forall y[[\forall x(Ax \blacktriangleright Bx) \wedge A(y)] \rightarrow B(y)]$ [where in $Ax \blacktriangleright Bx$, $x$ is not in the scope of any quantifier attached to the variable $y$]

**(II$_{\text{exp}}$):** $\vDash \forall x Bx \rightarrow\rightarrow \forall x(Ax \blacktriangleright Bx)$

**(III):** $\vDash [\forall x(Ax \blacktriangleright Bx) \wedge \forall x(Bx \blacktriangleright Cx)] \rightarrow \forall x(Ax \blacktriangleright Cx)$

**(III$_{\text{var}}$):** $[\forall x(Ax \blacktriangleright Bx) \wedge \neg \forall x(Ax \blacktriangleright Cx)] \rightarrow \neg \forall x(Bx \blacktriangleright Cx)$

**(IV$_{\text{exp}}$):** $\vDash \forall x(Ax \blacktriangleright Bx) \wedge \forall x(Ax \blacktriangleright Cx) \rightarrow\rightarrow \forall x(Ax \blacktriangleright (Bx \wedge Cx))$

**(IV$_{\text{var}}$):** $[[\forall x(Ax \blacktriangleright Bx) \wedge \neg \forall x[Ax \blacktriangleright (Bx \wedge Cx)]] \rightarrow \neg \forall x(Ax \blacktriangleright Cx)]$

**(V$_{\text{exp}}$):** $\vDash \neg \forall x(Ax \blacktriangleright Bx) \rightarrow\rightarrow \exists x(Ax \wedge \neg Bx)$

**(VI):** $\vDash \exists x(Ax \wedge \neg Bx) \rightarrow \neg \forall x(Ax \blacktriangleright Bx)$

**(C$_{\text{exp}}$):** $\vDash \forall x(Ax \blacktriangleright Bx) \Leftrightarrow \forall x(\neg Bx \blacktriangleright \neg Ax)$

The proofs are rather trivial given the results already established, but just to be explicit:

*Proof.* **($C_{exp}$)**: From our early list of strong equivalences we have $\vDash (Ax \blacktriangleright Bx) \Leftrightarrow (\neg Bx \blacktriangleright \neg Ax)$; by the Intersubstitutivity Corollary of Section 10, **($C_{exp}$)** follows.

**(I):** By L2b we have $\vDash [(Ay \blacktriangleright By) \wedge Ay] \rightarrow By$; and $\vDash \forall x(Ax \blacktriangleright Bx) \rightarrow (Ay \blacktriangleright By)$ since the antecedent strongly implies the consequent. So by L6$^{+}$, $\vDash [\forall x(Ax \blacktriangleright Bx) \wedge Ay] \rightarrow By$, which gives **(I)** by universal generalization (which is trivially valid, given the definition of validity and the quantifier rules).

**($II_{exp}$):** By L3a (since of course $\vDash B \Rightarrow (A \supset B)$) we have $\vDash Bx \rightarrow\!\!\rightarrow (Ax \blacktriangleright Bx)$. By generalizing and applying the Observation based on L11b, we get the result.

**(III):** By L6 we have $\vDash [(Ax \blacktriangleright Bx) \wedge (Bx \blacktriangleright Cx)] \rightarrow (Ax \blacktriangleright Cx)$. By generalizing and applying L11b, and using the strong equivalence of $\forall x[X \wedge Y]$ to $\forall x X \wedge \forall x Y$ and therefore their intersubstitutivity, we get **(III)**.

**($III_{var}$):** Applying L11b to the generalization of L7a, we get $\vDash \exists x[(A \blacktriangleright B) \wedge \neg(A \blacktriangleright C)] \rightarrow \exists x \neg(B \blacktriangleright C)$; but by one of the strong implications, also $\vDash [\forall x(A \blacktriangleright B) \wedge \exists x \neg((A \blacktriangleright C)] \rightarrow \exists x[(A \blacktriangleright B) \wedge \neg(A \blacktriangleright C)]$. By transitivity (and rewriting $\exists x \neg$ as $\neg \forall x$) we get the result.

**($IV_{exp}$):** By L5d, $\vDash (Ax \blacktriangleright Bx) \wedge (Ax \blacktriangleright Cx) \rightarrow\!\!\rightarrow (Ax \blacktriangleright (Bx \wedge Cx))$. Proceed as in **(III)**, but using L11b for one direction and L11c for the other.

**($IV_{var}$):** Applying L11b to the generalization of L7b-i, we get $\vDash \exists x[(A \blacktriangleright B) \wedge \neg(A \blacktriangleright C)] \rightarrow \exists x \neg(B \blacktriangleright C)$; from which we get the desired result by the method used for **($III_{var}$)**.

**($V_{exp}$):** By L3a we have $\vDash (Ax \supset Bx) \rightarrow\!\!\rightarrow (Ax \blacktriangleright Bx)$, so generalizing and using the Observation based on L11b, $\vDash \forall x(Ax \supset Bx) \rightarrow\!\!\rightarrow \forall x(Ax \blacktriangleright Bx)$; $\rightarrow\!\!\rightarrow$ contraposes, so this is equivalent to **($V_{exp}$)**.

**(VI):** By L4a we have $\vDash \neg(Ax \supset Bx) \rightarrow \neg(Ax \blacktriangleright Bx)$; so generalizing and using the $\exists$ half of L11b, $\vDash \exists x \neg(Ax \supset Bx) \rightarrow \exists x \neg(Ax \blacktriangleright Bx)$, which is equivalent to **(VI)**. □

I don't claim that every apparent law that might sound plausible has been established. In particular, certain plausible-sounding laws involving nested conditionals will fail.[29] Of course, any classically valid law of conditionals is bound to hold in contexts where the claims involved obey strong excluded middle, since all the conditionals reduce to $\supset$ there; but it might seem disappointing that they don't hold more generally. But I remind the reader of a remark made at the end of Section 1: the goal wasn't to validate every law that superficially "sounds right", it was to validate enough to yield a logic with which we could actually reason. And I hope that the laws I've established at least give initial plausibility to the idea that with the conditionals here or something close to them, we could do that.

**§12. The prospects for other approaches.** I expect that some variations on the approach to restricted quantification presented in this paper are possible within naive truth

[29] Some failures, such as "If $x$ is $A$, then if all $A$ are $B$ then $x$ is $B$", can be handled by the special convention that in English, "If $A$ then if $B$ then $C$" is usually read as meaning "If $A$ and $B$ then $C$". (There seems to be no natural such convention that would have avoided the problem posed by restricted quantification.)

theory;[30] indeed, I'd expect, and hope for, variations that give rise to clear improvements in the laws. But what about the prospects for a radically different approach?

One thing one can certainly hope for is an approach involving only a single conditional. I've remarked that this inevitably requires the failure of some of the laws on my menu from Section 1 (and the expanded menu in Section 3), but one might be able to swallow that if suitably strong approximations to them were available. One natural idea is to look for a naive truth theory for the logic CK (also called RWK), which is a weakening of Lukasiewicz continuum-valued logic; I won't specify it in detail (you can find it in Priest, 2008, p. 228), but it is a nonparaconsistent logic, and the laws involving the conditional include modus ponens plus $A \to A$ plus

$$\begin{aligned}
&\textbf{(CONTRAP):}\ \vDash (A \to B) \to (\neg B \to \neg A)\\
&\textbf{(CONJ):}\ \vDash [(A \to B) \wedge (A \to C)] \to [A \to (B \wedge C)]\\
&\textbf{(WEAKENING):}\ \vDash B \to (A \to B)\\
&\textbf{(ASSERTION):}\ \vDash A \to [(A \to B) \to B]\\
&\textbf{(SUFFIX):}\ \vDash (A \to B) \to [(B \to C) \to (A \to C)].
\end{aligned}$$

(Note that while there are analogs of all of these in the system I've given, most require the two distinct conditionals and hence are weaker.) It is not known whether a naive truth theory is possible in CK—more on this in a moment—but even if it is, one must give up on items **(I)**, **(III)**, **(III$_{\mathbf{var}}$)**, **(IV$_{\mathbf{var}}$)**, and **(VI)** from my initial menu. **(I)** must go because as remarked earlier, it leads to Curry's paradox in a single-conditional theory. **(VI)** must go because with even the rule form of contraposition and $A \to A$ it requires excluded middle. **(III$_{\mathbf{var}}$)** and **(IV$_{\mathbf{var}}$)** must go because each leads to excluded middle for conditional

[30] One variation that will be obvious to those familiar with Field (2008) is to use revision-theoretic constructions for the fibers and/or the base-space, modeled after the fixed point constructions used here. For instance, to construct the fiber for a given $\to$-booster $g$, we'd let

$$v_{g,\mu}(A \rhd B) = \begin{cases} 1 & \text{if } (\exists\kappa < \mu)(\forall\xi \in [\kappa, \mu))(\text{if } |A|_{g,\xi} = 1 \text{ then } |B|_{g,\xi} = 1) \\ 0 & \text{if } (\exists\kappa < \mu)(\forall\xi \in [\kappa, \mu))(|A|_{g,\xi} = 1 \wedge |B|_{g,\xi} = 0) \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

The fiber for $g$ would consist of the members of this sequence that keep coming back; its base node would be $v_{g,\Delta}$ where $\Delta$ is a reflection ordinal for the sequence (in the sense of Gupta & Belnap, 2003, p. 172). This would yield a Regularity Theorem for fibers. Similarly, letting $||A||_g$ be the value of $A$ at the reflection ordinal for $g$, we'd let

$$g_\alpha(A \to B) = \begin{cases} 1 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta, \alpha))(\text{if } ||A||_{g_\gamma} = 1 \text{ then } ||B||_{g_\gamma} = 1) \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

The base space would consist of the members of this sequence that keep coming back, and the principal node would be the value at the reflection ordinal for this sequence. (The Kripke construction would be like the one in the present paper: an $\to$ -conditional can get value 0 even when its $g$-value is $1/2$.) The "Fundamental Theorem" of Field (2008) would fail for the same reason that "Regularity" fails for the base space in the present construction.

I haven't checked this out fully, but I doubt that it would lead to very different laws than those given by the present construction. I chose the fixed point construction rather than the revision construction in this paper partly because it is new and partly because it strikes me as somewhat more natural.

sentences,[31] and CK is incompatible with that since it requires the Curry sentence to behave like the Liar sentence. **(III)** must go because it requires

**(CS):** $[(A \to B) \wedge (B \to C)] \to [A \to C]$

("single conditional conjunctive syllogism") which, given (ASSERTION) and (SUFFIX), leads to **(I)** and thus the Curry paradox.[32]

But what about "strong approximations" to **(I)**, **(III)**, **(III$_{\mathbf{var}}$)**, **(IV$_{\mathbf{var}}$)**, and **(VI)**? Because of (ASSERTION), we would get

**(I-alt):** $\vDash$ If $x$ is $A$, then if all $A$ are $B$ then $x$ is$B$.

Also, it's a feature of CK that we can define in it a strengthened conjunction $\circ$ called "fusion", with many of the properties of conjunction: $A \circ B$ is $\neg(A \to \neg B)$. Perhaps a sufficiently strong approximations to **(I)**, **(III)**, **(III$_{\mathbf{var}}$)**, and **(IV$_{\mathbf{var}}$)** are **(I-alt)** and

**(I-fus):** $\vDash$ If (all $A$ are $B$)$\circ$ ($x$ is $A$) then $x$ is $B$,

**(III-fus):** $\vDash$ If (all $A$ are $B$)$\circ$ (all $B$ are $C$) then all $A$ are $C$

**(III$_{\mathbf{var}}$-fus):** $\vDash$ If (all $A$ are $B$)$\circ$ (not all $A$ are $C$) then not all $B$ are $C$

**(IV$_{\mathbf{var}}$-fus):** $\vDash$ If (all $A$ are $B$)$\circ$ (not all $A$ are both $B$ and $C$) then not all $A$ are $C$.

And perhaps a sufficiently strong approximation to **(VI)** is its rule form

**(VI-w):** Something is both $A$ and not $B \vDash$ Not all $A$ are $B$.

These are consequences of CK, so if they are satisfactory surrogates then we could handle restricted quantification in a naive truth theory with a single conditional by getting a naive truth theory for CK.

But as I've said, no one knows how to give a naive truth theory for CK—the question of whether this could be done was explicitly raised in print 20 years ago (Restall, 1992), but no one seems to have made significant progress on it.[33] Probably the naive truth theories

---

[31] It suffices to show that both lead to $[(A \to B) \wedge \neg(A \to B)] \to \perp$, since by (CONTRAP) and modus ponens and Kleene laws that yields $(A \to B) \vee \neg(A \to B)$. But ($\text{III}_{\text{var}}$) requires $[(A \to B) \wedge \neg(A \to C)] \to \neg(B \to C)$, and if one takes $C$ to be $B$ one effectively gets $[(A \to B) \wedge \neg(A \to B)] \to \perp$. And ($\text{IV}_{\text{var}}$) requires $[(A \to B) \wedge \neg[A \to (B \wedge C)]] \to \neg(A \to C)$, and if one takes $C$ to be $A$ one effectively gets the same thing.

[32] Greg Restall pointed this out to me some years back. Argument sketch:

(1) A special case of (ASSERTION) is $B \to [(B \to B) \to B]$; and it's easy to get from that to $[B \wedge (B \to C)] \to [[(B \to B) \to B] \wedge (B \to C)]$.

(2) A special case of (CS) is $[[(B \to B) \to B] \wedge (B \to C)] \to [(B \to B) \to C]$.

(3) From these together we get $[B \wedge (B \to C)] \to [(B \to B) \to C]$, by transitivity.

(4) In Anderson & Belnap (1975, pp. 79–80) it is observed that (ASSERTION) and (SUFFIX) suffice for (PERM), that is, the schema $\vDash [A \to (B \to C)] \to [(B \to (A \to C)]$. (Use the instance $B \to [(B \to C) \to C]$ of (ASSERTION), and apply the rule form of (SUFFIX) to get $[[(B \to C) \to C] \to (A \to C)] \to [B \to (A \to C)]$. So to get to (PERM) we need only $[A \to (B \to C)] \to [[(B \to C) \to C] \to (A \to C)]$, and that is an instance of (SUFFIX).)

(5) But applying (PERM) to the result of (3), we get $(B \to B) \to [[B \wedge (B \to C)] \to C]$. So by $B \to B$ and modus ponens we derive pseudo-modus ponens, that is, **(I)**.

[33] I have heard claims in conversation that it is known to be possible, but they rest on at least one of the following weakenings of the demands. One weakening is to a logic without the ordinary conjunction and disjunction. The other weakening gives up the requirement of conservativeness in the sense of Section 4, replacing it with a mere consistency requirement. That somewhat trivializes

in print for logics that come *closest* to CK (though they don't come very close at all) are my own in (2008) and several much earlier ones of Ross Brady's, most of which are discussed in his (2006). Neither he nor I claimed to give an account of account of restricted quantification; and if one had defined the restricted universal quantifier in terms of the conditionals he or I provided, the accounts would have been ludicrous.

In the case of Field (2008), the logic contains (CONTRAP), but only the rule forms of (CONJ), (WEAKENING)[34], and (SUFFIX), and not even the rule form of (ASSERTION), that is, not even

**(WEAK ASSERTION):** $A \vDash (A \to B) \to B$.

The laws of restricted quantification that would have been delivered would have been, to a close approximation, those that would be delivered by the construction of this paper *were all "if...then's" understood as* ►. This means that not only **(I)** would fail (as is inevitable on a single-conditional construction), but so would **(I-alt)**, and indeed so would even

**(I-vw):** $x$ is $A$ $\vDash$ If all $A$ are $B$ then $x$ is $B$

and

**(I⁻-vw):** Everything is $A$ $\vDash$ If all $A$ are $B$ then everything is $B$.

**(II)** and **(V)** would fail too: only their rule forms would hold (though as regards **(II)**, see note 34). **(III)** would fail too: the best we could get is

**(III-vw):** All $A$ are $B$ $\vDash$ If all $B$ are $C$ then all $A$ are $C$

and the analog with 'All $A$ are $B$' and 'All $B$ are $C$' switched. Analogously for **(III$_{\mathbf{var}}$)**, **(IV)**, and **(IV$_{\mathbf{var}}$)**. (And **(VI)** would also hold only in weakened rule form, but that's inevitable on a single conditional theory). As I said, that would be ludicrous for an account of restricted quantification. Brady's preferred one-conditional logics are in some respects stronger than the one in Field (2008): they contain (CS), and the full (CONJ) rather than just the rule form. But like Field (2008) they don't validate even (WEAK ASSERTION), so were one to define restricted quantification one wouldn't get even the very weak forms **(I-vw)** and **(I⁻-vw)** of **(I)** and **(I⁻)**; and as with Field (2008) one would get only the rule form of **(VI)**.

But things would be much worse: it's a feature of the Brady semantics that conditionals are often evaluated as false when they have true consequents or false antecedents. Because of this, if one used this conditional for restricted quantification not even the rule forms of **(II)** or **(V)** would hold. Thus from "Everything is $B$" one wouldn't be able to infer "All $A$ are $B$", and from "Not all $A$ are $B$" one wouldn't be able to infer "Something is both $A$ and not $B$". That would be even more ludicrous as an account of restricted quantification than using Field (2008). For these reasons, Brady prefers a two-conditional account of restricted quantifiers (Beall *et al.* 2006), to be discussed below.

---

the task at hand, since it's been known for many years that a naive truth theory is consistent even in Lukasiewicz continuum valued logic (see White, 1979); the problem is that in that logic the theory isn't $\omega$-consistent, that is, consistent with a standard model of syntax (see Restall, 1992), which also prevents a consistent naive satisfaction theory with the usual compositional rules for the quantifiers (see Hajek *et al.*, 2000). It was partly to prevent such pathologies that I insisted on a conservativeness requirement.

34 Though I only realized it later, the trivial way of getting (WEAKENING) while retaining (CONTRAP), namely defining a new conditional in terms of the old by disjoining with the material conditional, seems not to have a cost with respect to other laws of this system. Though that's largely because the other laws are so weak!

Part of the problem with using either the logic of Field (2008) or any of the main logics of Brady (2006) to define restricted quantification is that in them, the single-conditional is contraposable (as it is in CK). If one is intent on a single-conditional approach, one might do better exploring logics where it isn't (even in rule form). Andrew Bacon (2013) has recently made considerable progress along these lines: he uses a novel construction based on the Banach Fixed Point Theorem. (He presents the construction for a language without a real negation, but has suggestions for how such a negation could be added; when I say that he's concerned with a noncontraposable conditional I mean that it would be noncontraposable were negation added.[35]) His construction validates many theorems of (CK) and some nontheorems such as (CS); this last is nice because it gives **(III)**. The main downside is that not only does it not yield **(I)** (as is inevitable in a single-conditional theory with modus ponens), it also doesn't yield **(I-alt)** or even **(I-vw)** or **($\mathbf{I^-}$-vw)**.[36] Also, once negation is added it not only doesn't yield **($\mathbf{I_c}$)**, it doesn't yield even the rule forms of **($\mathbf{II_c}$)** and **(V*)** either. I think any of these makes it inadequate as a single-conditional theory. On the other hand, maybe it or something like it could serve as one conditional in a two-conditional theory? That's definitely worth thinking about. (I only saw Bacon's paper after I'd completed this one, so I'm going to have to leave matters at that.)

One could also explore the possibility of two-conditional theories where the conditional with respect to which restricted quantification is defined is just $(A \rightarrow B) \wedge (\neg B \rightarrow \neg A)$, where $\rightarrow$ is the conditional used for 'if ... then'. I know of no reason in principle why one couldn't get all the target laws including (I) in such a theory, and if one could it would have obvious advantages of theoretical economy over the approach I've given. But I have no idea how to give a conservativeness proof for naive truth in any such theory.

I turn last to attempts within the paraconsistent tradition for naive truth theory. Several authors in this tradition appreciated the problem posed by restricted quantification well before I did. The two main discussions I know of are Beall *et al.* (2006) and Beall (2009). In both cases, the authors first consider single conditional accounts and find them wanting. Their reason is that the single conditional '$\rightarrow$' that they consider, like the '$\rightarrow$' in this paper, does not satisfy the rule

**(WEAK$\rightarrow$-WEAKENING):** $B \vDash A \rightarrow B$.

So they propose two-conditional accounts, where the second conditional (which I'll write as $\triangleright$, though neither it nor the $\rightarrow$ are the same as the ones in prior sections of this paper) does obey the analogous

**(WEAK $\triangleright$-WEAKENING):** $B \vDash A \triangleright B$.

[35] Of course as he notes, one can easily define a contraposable conditional from a noncontraposable one; but the new one may not validate the laws that made the old one attractive.

[36] Bacon has since pointed out to me that this is inevitable in a one-conditional naive theory that validates (WEAKENING), (CONJ), and (CS) (and modus ponens and very weak additional assumptions). The argument, which he credits to Tore Fjetland Ogaard, uses the Curry sentence $K$. (WEAKENING) gives $K \rightarrow (\top \rightarrow K)$, and the defining property of $K$ (plus naivety) gives $K \rightarrow (K \rightarrow \bot)$; (CONJ) then yields $K \rightarrow [(\top \rightarrow K) \wedge (K \rightarrow \bot)]$. And (CS) gives $[(\top \rightarrow K) \wedge (K \rightarrow \bot)] \rightarrow (\top \rightarrow \bot)$. By a very weak transitivity rule (which follows from (CS), modus ponens and $\wedge$-Introduction), the last two conclusions give $K \rightarrow (\top \rightarrow \bot)$. But **(I-vw)** is effectively the same as (WEAK ASSERTION), which gives $(\top \rightarrow \bot) \rightarrow \bot$; so using the weak transitivity rule again we get $K \rightarrow \bot$. That's equivalent to $K$, so modus ponens yields triviality. I think this argument shows that (CS) has an unexpectedly high cost in a single-conditional theory, even without contraposition.

(I use $\rhd$ rather than $\blacktriangleright$ for their second conditional since they take it to be noncontraposable.) So far very much like the account in this paper. However, though the second conditional in Beall *et al.* (2006) obeys the $\rhd$-weakening *rule* and hence delivers *the rule form of* **(II)**, it doesn't satisfy the stronger $\rhd$-weakening axiom and so does not deliver the full **(II)**.

This seems a substantial limitation. Beall (2009) recognized this, and his preferred account (p. 125) gets the strengthening by a different version of $\rhd$ from that in Beall *et al.* (2006): he uses a $\rhd^*$, with $A \rhd^* B$ defined to be simply $(A \rightarrow B) \vee B$, where $\rightarrow$ is the conditional of the 2006 paper. But though this gives **(II)** in full form, it has tremendous costs: in particular, it precludes the validity of even the rule form of CONJ (and hence even the rule form of **(IV)**). The problem is that if $A \rhd^* B$ holds by virtue of $A \rightarrow B$ holding, and $A \rhd^* C$ holds by virtue of $C$ holding, then one can't infer $A \rhd^* (B \wedge C)$ unless one can infer it from $(A \rightarrow B) \wedge C$; and since with the 2006 conditional Beall is working with one can't infer $A \rightarrow C$ from $C$, there is no way to do this.

This problem would be alleviated had he started not from the $\rightarrow$ in Beall *et al* (2006) but from the $\rhd$ in that paper (or some other conditional for which (WEAK WEAKENING) holds), that is, had he defined $A \rhd^* B$ as $(A \rhd B) \vee B$;[37] but even then, the move would mean that even if the full (CONJ) holds for $\rhd$, only the rule form holds for $\rhd^*$. In short, using the disjunctive trick to strengthen **(II)** from a rule to the full conditional results *at best* in a corresponding weakening of **(IV)**.

Another matter: in both Beall *et al.* (2006) and Beall (2009) it is assumed that the conditional $\rhd$ that goes with the restricted quantifier is to be "rule-weaker than" $\rightarrow$, in the sense that

**(W):** $A \rightarrow B \vDash A \rhd B.$

This imposes a high cost: it rules out **(I)**. (For with (W), **(I)** requires single-conditional pseudo-modus-ponens for $\rhd$, which is incompatible with naive truth.) Indeed, those papers don't even deliver **(I-vw)** and **(I⁻-vw)**, though I see no obvious reason why some other paraconsistent account couldn't have these even if it contained (W).

To partially summarize what we have so far, no available paraconsistent two-conditional account delivers even **(I-vw)** or **(I⁻-vw)**, and none delivers both the full **(II)** and the full **(IV)**. But as far as I know there is no reason why improved paraconsistent theories couldn't obtain these.

But there are also in principle limitations on what any approach to the paradoxes that admits acceptance gluts (and hence disallows the explosion rule $A, \neg A \vDash B$) could deliver. The chief such limitation is that (assuming that the restricted quantifier conditional $\rhd$ obeys modus ponens) we cannot have the rule

$$\neg A \vDash A \rhd B,$$

since with modus ponens it leads to explosion. It follows that we can't have

$$\neg(A \wedge \neg B) \vDash A \rhd B,$$

since its antecedent is weaker. Because of these, we can't have *even the rule forms of* **($\mathbf{II_c}$)** (from Section 1) or **(V*)** (from Section 3): for instance, we can't infer from "Nothing is

[37] He didn't because the account in Beall *et al* (2006) hasn't actually been shown consistent with naive truth theory, it is merely conjectured to be.

both $A$ and not $B$" to "All $A$ are $B$".[38] (Or rather: a paraconsistent dialetheist can keep these inferences only by disallowing modus ponens for $\rhd$, and thus disallowing the inference from "All $A$ are $B$" and "$c$ is $A$" to "$c$ is $B$".) These limitations are fairly obvious, and presumably anyone with an independent commitment to accepting both the Liar sentence and its negation will be prepared to swallow them. But they strike me as a heavy intuitive cost, and this seems to me to give a reason to prefer rejecting both the Liar sentence and its negation rather than accepting them.

**§13. Acknowledgments.** Thanks to two anonymous referees for helpful suggestions that have led to improvements.

## BIBLIOGRAPHY

Anderson, A. R., & Belnap, N. (1975). *Entailment*, Vol. 1. Princeton, NJ: Princeton University Press.

Bacon, A. (2013). A new conditional for naive truth theory. *Notre Dame Journal of Formal Logic*, **54**, 87–103.

Beall, J. C. (2009). *Spandrels of Truth*. Oxford, UK: Oxford University Press.

Beall, J. C., Brady, R., Hazen, A., Priest, G., & Restall, G. (2006). Relevant restricted quantification. *Journal of Philosophical Logic*, **35**, 303–324.

Brady, R. (2006). *Universal Logic*. Stanford, CA: CSLI Publications.

Field, H. (2003). No fact of the matter. *Australasian Journal of Philosophy*, **81**, 457–480.

Field, H. (2004). The semantic paradoxes and the paradoxes of vagueness. In Beall, J. C., editor. *Liars and Heaps*. Oxford, UK: Oxford University Press, pp. 262–311.

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

Gupta, A., & Belnap, N. (2003). *The Revision Theory of Truth*. Cambridge, MA: MIT Press.

Hajek, P., Paris, J., & Sheperdson, J. (2000). The liar paradox and fuzzy logic. *Journal of Symbolic Logic*, **65**, 339–346.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge, UK: Cambridge University Press.

Horsten, L. (2011). *The Tarskian Turn: Deflationism and Axiomatic Truth*. Cambridge, MA: MIT Press.

Priest, G. (2008). *An Introduction to Non-classical Logic* (second edition). Cambridge, UK: Cambridge University Press.

Restall, G. (1992). Arithmetic and truth in Lukasiewicz's infinitely-valued logic. *Logique et Analyse*, **139–140**, 303–312.

White, R. (1979). The consistency of the axiom of comprehension in the infinite-valued predicate logic of Lukasiewicz. *Journal of Philosophical Logic*, **8**, 509–534.

DEPARTMENT OF PHILOSOPHY, NEW YORK UNIVERSITY
5 WASHINGTON PLACE, NEW YORK NY 10003, USA
and

DEPARTMENT OF PHILOSOPHY, UNIVERSITY OF BIRMINGHAM
EDGBASTON, BIRMINGHAM, B15 2TT UK
*E-mail*: hartry.field@nyu.edu

---

[38] There is no obvious bar to a paraconsistentist accepting **(II)** and **(V)**, as long as she denies the contraposibility of $\rhd$.