THE REVIEW OF SYMBOLIC LOGIC
Volume 0, Number 0, xxxx 2015

# INDICATIVE CONDITIONALS, RESTRICTED QUANTIFICATION, AND NAIVE TRUTH

HARTRY FIELD

**Abstract.** This paper extends Kripke's theory of truth to a language with a variably strict conditional operator, of the kind that Stalnaker and others have used to represent ordinary indicative conditionals of English. It then shows how to combine this with a different and independently motivated conditional operator, to get a substantial logic of restricted quantification within naive truth theory.

**§1. Introduction.** The "naive" notion of truth, according to which for each sentence $S$ of our language, the claim that $S$ is true is equivalent to $S$ itself,[1] appears at first blush to be doomed by the Liar paradox and other related paradoxes. But only at first blush: one of the lessons that can be drawn from Kripke 1975 is that naivety in a theory of truth can be retained if one is willing to give up the hegemony of classical logic. There is little reason to doubt the correctness of classical logic as applied to our most serious discourse, e.g. our most serious physical theories. But the semantic paradoxes arise because truth talk gives rise to some anomalous applications (e.g. "viciously self-referential" ones), and it's rash to assume that classical logic continues to be appropriate to these applications. Maybe we should generalize logic in a way that allows these anomalies to be treated non-classically, while enforcing classicality in situations where anomalies can't arise. Kripke's paper, in particular the parts concerning logics based on Kleene valuation schemes, suggests the possibility of naive truth in this setting: in particular, one can have naive truth in a logic that restricts the general application of excluded middle, but which reduces to classical logic in contexts where the anomalies of truth cannot occur.

It isn't *immediately obvious* that the best response to the paradoxes is to abandon the hegemony of classical logic while retaining the hegemony of naive truth—*prima facie*, the reverse seems at least as attractive. But the costs of restricting naive truth turn out to be extraordinarily high,[2] and so the program of trying to keep it by restricting the scope of classical logic is one well worth pursuing. Kripke 1975 was the first substantial step.[3]

Kripke's paper by itself shows the possibility of naive truth only for languages of very limited expressive power. The question arises as to how far his ideas can be generalized,

Received: April 27, 2015.

[1] I ignore ambiguities, indexical elements, etc., so as to be able to concentrate on sentence-*types*. There are subtleties about how best to extend the idea of naive truth to token utterances, but I will not be concerned with those issues here.

[2] See Field 2008, Part II, for a review.

[3] In Kripke's paper, and in the present paper too, we keep the classical *structural* rules for validity: (a) validity is transitive (in the general form given by the Cut Rule), and (b) valid inference is a relation between a *set* of premises and a conclusion (as opposed e.g. to a multi-set, where the number of occurrences of the premise matter, as in logics without structural contraction). The use of substructural logics is unnecessary.

There is also no need to restrict reasoning by cases, or to embrace dialetheism.





doi:10.1017/S1755020315000301

ZU064-05-FPR S1755020315000301 19 September 2015 12:42



and on this there has been some progress in recent years. In particular, there are now techniques for generalizing it to include certain kinds of conditionals (despite the threat of Curry-like paradoxes)[4]. But one kind of conditional operator that has not been treated in the literature on naive truth is "variably strict" conditional operators of the sort that have been discussed by Stalnaker 1968, Lewis 1974, Pollock 1976, Burgess 1981, and many others. The rough idea of such a conditional is that it is true at a world $w$ if and only if at all worlds $x$ *where its antecedent is true but that are otherwise only minimally different from* $w$, its consequent is true. (There are different ways of spelling out this rough idea, depending mostly on the assumptions made about a relation of relative closeness of worlds; in this paper I'll adopt a framework, Burgess semantics, that is as neutral as possible about this.) Variably strict conditionals are clearly non-monotonic ('If $A$ then $C$' doesn't imply 'If $A$ and $B$ then $C$'); from which it pretty much follows that they are non-transitive.[5] (They are also non-contraposable.) Their non-monotonicity and resulting non-transitivity make them significantly different from the sort of conditionals heretofore discussed in the naive truth literature. The early parts of the present paper provide a method (actually more than one) of extending Kripke's theory to cover languages with such a variably strict conditional—including in Section 6 the important case of languages that also have another conditional operator for restricted quantification.

Proponents of variably strict conditionals have divided over how extensive their application is. Some, e.g. Lewis, have taken a variably strict operator to model only "counterfactual" or "subjunctive" conditionals of ordinary language, and have held that "indicative conditionals" of ordinary language are represented by the familiar '$\supset$'. But it's well known that understanding ordinary indicatives in terms of '$\supset$' is *prima facie* counterintuitive—e.g. on that understanding, "If I run for President, I'll be elected" comes out true, since I'm resisting all pressure to run—and nowadays it's more common to think, with Stalnaker, that the variably strict conditional account is applicable to ordinary indicative conditionals as well as "counterfactuals". The first six sections of this paper are neutral on this issue.

But I favor the Stalnaker position, and this is relevant to an important application of the material in the early sections to the logic of restricted quantification, in Section 7. Restricted quantification poses a serious challenge to naive truth theory. In such a theory there are already difficulties with properly handling ordinary restricted quantifications like "Every true sentence in Jones' book appeared earlier in Smith's", but the difficulties become far greater when one tries to come up with a plausible account of how these interact with conditionals in a way that validates plausible laws such as "If all $A$ are $B$ and $y$ is $A$ then $y$ is $B$" and "If everything is $B$ then all $A$ are $B$". I've addressed this challenge before (Field 2014), but in a rather *ad hoc* manner; an ultimate goal of this paper is to answer the challenge without *ad hoc*ness, by bringing in a more general logic of indicative conditionals.

**§2. Two-valued and three-valued worlds models for the language of indicative conditionals.** Let $L$ be a language whose logical primitives are '$\neg$', '$\wedge$', '$\forall$', '$=$', a unary necessity operator '$\Box$', and a binary conditional operator '$\triangleright$'. An additional conditional for restricted quantification will be added in Sections 5 and 6. For the moment, let's suppose that $L$ doesn't contain "paradox-prone" terms like 'True' that will require special treatment.

---

[4] See Restall 2007 for a discussion of such paradoxes and of the difficulties that a naive truth theory must overcome if it is to handle them.

[5] 'If $A$ and $B$ then $A$' is clearly valid for them, and with it, transitivity would imply monotonicity.

'$\rhd$' is supposed to represent the indicative and/or counterfactual conditional of English and be a "variably strict" conditional in the general ballpark of Lewis, Stalnaker, Pollock and Burgess. Of these semantics, Burgess's is the most general (that is, the others can be obtained by adding restrictions to it),[6] and I will consider both it and a slight modification of it. Both versions of the Burgess semantics are initially based on "2-valued worlds models", which I'll now describe. (For simplicity I'll assume that $L$ has no individual constants or function symbols; also, that its only variables are first order.)

A *2-valued worlds model* $M$ for $L$ consists of

**(i):** A non-empty set $W$ of worlds, perhaps with a distinguished non-empty subset $NORM$ of "normal" worlds. (Nothing central to this paper depends on allowing non-normal worlds; I do so simply for added generality. The definition of validity will be in terms of the normal worlds only, but allowing for non-normal worlds may affect which conditionals can be true at normal worlds.)

**(ii):** For each $w \in W$, a subset $W_w$ of $W$ and a pre-order (reflexive and transitive relation) $\leq_w$ on $W_w$.[7] (Think of $W_w$ as the set of worlds "accessible from" $w$, and '$x \leq_w y$' as meaning "the change from $w$ to $x$ is no greater than the change from $w$ to $y$".)

**(iii):** For each $w \in W$, a non-empty set $U_w$ (the universe of $w$). Let $U$ be the union of the $U_w$.

**(iv):** For each $w \in W$ and $k$-place predicate $p$, a function $p_w$ from $U^k$ (the set of $k$-tuples of members of $U$) to $\{0, 1\}$. (The set of $k$-tuples that get assigned value 1 is the *extension of* $p$ in the model.) We require that the function $=_w$ (associated with '=') assigns 1 to $< o, o >$ for each $o \in U$ and assigns 0 to all other pairs.

($W$, $NORM$, etc. can all vary from one model to another, so we should really write $W_M$, $NORM_M$, $W_{M,w}$, $\leq_{M,w}$, $U_{M,w}$ and $p_{M,w}$.) Regarding (iv), we could if we like impose the ("actualist") requirement that $p_w$ never assign value 1 to $k$-tuples not in ${U_w}^k$; it won't matter for what follows.[8]

Regarding (ii), we could if we like impose additional conditions on $W_w$ and $\leq_w$ for each $w \in W$, or at least for each $w$ in $NORM$. (The distinction of non-normal worlds from normal ones only matters if some such additional conditions apply only to normal worlds.) Indeed, one such condition is almost universally regarded as appropriate for indicative and counterfactual conditionals (at least for worlds $w$ in $NORM$):

**Weak Centering:** $w \in W_w$, and for any $x$ in $W_w$, $w \leq_w x$

That Weak Centering holds at least for worlds in $NORM$ is required if Modus Ponens for $\rhd$ is to be valid, on the account of validity soon to be given, which involves preservation of

[6] Not every defensible model of conditionals can be fit into the Burgess framework (or the slight modification of it to be mentioned soon). I suspect that the basic ideas of this paper can be adapted to plausible alternative models, but will not attempt to prove this.

[7] An alternative convention is to take $\leq_w$ to be a pre-order on the full $W$, and subject to the constraint that if $y \in W_w$ and $x \leq_w y$ then $x \in W_w$.

[8] In the 3-valued context to be introduced shortly, we could introduce a more thorough actualism, in which the $p_w$ never assign value 0 or 1 to such $k$-tuples; in effect this would make ${U_w}^k$ rather than the full $U^k$ the domain of $p_w$. But again, this would make no difference to the issues I'm concerned with.

value 1 at normal worlds.[9] (Modus Ponens has been questioned for indicative conditionals (McGee 1985), but the grounds for doing so seem weak in the context of the semantics for variably-strict conditionals.)[10]

In addition to Weak Centering, Lewis, Stalnaker, Pollock and many others also accept one or more of the following conditions (for all worlds or just for normal ones):

**Strong Centering:** $w \in W_w$, and for any $x$ in $W_w$ other than $w$, $w <_w x$ (i.e. $w \leq_w x$ and not($x \leq_w w$))
**No Incomparabilities:** for any $x$, $y$ in $W_w$, either $x \leq_w y$ or $y \leq_w x$
**No Ties:** for any distinct $x$, $y$ in $W_w$, not both $x \leq_w y$ and $y \leq_w x$
**Limit Condition:** the relation $<_w$ is well-founded.

What follows will be completely neutral as to which if any such conditions are imposed, except for occasional reminders that restricting to models with Weak Centering (at least at normal worlds) is advantageous.[11]

To simplify the presentation of the semantics I adopt the usual trick of expanding the language to contain a new name for each object in $U$; call the expanded language $L^+$. (The expansion depends on the underlying model, so we should really write $L^+_M$.) I'll consider two ways of evaluating the sentences of $L^+$ in $M$.

The first version is 2-valued:

**Burgess evaluation procedure:**

- $|p(c_1, ..., c_k)|_w$ is just $p_w(o_1, ...o_k)$, where $c_1, ..., c_k$ are the names for $o_1, ..., o_k$ respectively.

---

[9] Demanding Weak Centering at non-normal worlds as well as normal ones would lead in addition, in the current 2-valued framework, to the validity of the inference from $C \rhd A$ and $C \rhd (A \rhd B)$ to $C \rhd B$. If we want Modus Ponens without getting that even for 2-valued sentences, we need the flexibility provided by non-normal worlds. In general, the point of non-normal worlds is to provide such added flexibility as to what comes out valid.

I've said that nothing in this paper depends on making use of such added flexibility: there will be no need to have the flexibility in the logic that includes 'True' if one doesn't utilize it in the base logic without 'True'. This may seem surprising: we presumably want Modus Ponens for $\rhd$, but we don't want the law just cited since by taking $A$ to be $C$ we'll be led to the inference from $C \rhd (C \rhd B)$ to $C \rhd B$, which in combination with Modus Ponens is well known to rule out naive truth by Curry's paradox. But there is actually no problem: in the semantics to be introduced, Weak Centering at *all* worlds guarantees only that the inference from $C \rhd A$ and $C \rhd (A \rhd B)$ to $C \rhd B$ will hold *for 2-valued sentences*; and the sentences involved in Curry-type paradoxes will not be 2-valued. (Modus Ponens, on the other hand, will be guaranteed for *all* sentences, even by Weak Centering just at normal worlds.)

[10] The canonical supposed counterexample involves a 3-candidate race whose leading candidates are a Democrat and a Republican, with an Independent far behind. Then the claim "If the Republican doesn't win, the Independent will" seems false. But "The Democrat won't win" may be true, and "If the Democrat doesn't win, then if the Republican doesn't win the Independent will win" may seem true; and these two claims lead to the false claim by Modus Ponens. A standard resolution of this, which I support, is that the complex conditional that "may seem true" isn't: what's true is only that if the Democrat doesn't win *and* the Republican doesn't win then the Independent will win, but to get from that to the complex conditional one needs the rule of Exportation $(A \wedge B) \rhd C \models A \rhd (B \rhd C)$, which is invalid on the variably-strict semantics.

[11] It's also possible to add "purely modal" conditions, not involving the $\leq_w$; e.g.

**S4** if $x \in W_w$ and $y \in W_x$ then $y \in W_w$.

What follows is neutral on these as well.

- $|\neg A|_w$ is $1-|A|_w$
- $|A \wedge B|_w$ is $\min\{|A|_w, |B|_w\}$
- $|\forall x A|_w$ is $\min\{|A(c/x)|_w$ : all $c$ that name members of $U_w\}$
- $|\Box A|_w$ is $\min\{|A|_x : x \in W_w\}$
- $|A \rhd B|_w = \begin{cases} 1 & \text{iff } (\forall x \in W_w)[|A|_x = 1 \supset \\ & (\exists y \leq_w x)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 1)]] \\ 0 & \text{iff } (\exists x \in W_w)[|A|_x = 1 \wedge \\ & (\forall y \leq_w x)[|A|_y = 1 \supset (\exists z \leq_w y)(|A|_z = 1 \wedge |B|_z = 0)]] \end{cases}$

(Let a *w-neighborhood* be a non-empty subset $N$ of $W_w$ such that if $x \in N$ and $y \leq_w x$ then $y \in N$; and call a $w$-neighborhood *A-consistent* if it contains a world where $|A|$ is 1. Then the right hand side of the 1-clause for $\rhd$ says that all $A$-consistent $w$-neighborhoods have $A$-consistent sub-$w$-neighborhoods throughout which if $|A|$ is 1, so is $|B|$; and the right hand side of the 0-clause says that there is an $A$-consistent $w$-neighborhood such that every $A$-consistent sub-$w$-neighborhood of it contains a world where $|A|$ is 1 and $|B|$ is 0. If one were to make the "No Incomparabilities" assumption (for all worlds, not just normal ones) one could simplify these clauses for $\rhd$ a bit: that assumption amounts to the assumption that for each $w$, the $w$-neighborhoods are nested; and given that, the 1-clause is equivalent to the claim that if there is at least one $A$-consistent $w$-neighborhood then there is one throughout which if $|A|$ is 1, so is $|B|$.)

These stipulations give every $L^+$-sentence a unique value in $\{0,1\}$ at each world, given any 2-valued worlds model $M$. Conditionals don't in general contrapose, but they shouldn't: 'If Trump runs for President he won't be elected' shouldn't imply 'If Trump is elected he won't have run'.

Validity is explained as follows:

**(VAL):** An inference from a set $\Gamma$ of $L$-sentences to an $L$-sentence $B$ is *Burgess-valid* if for every worlds model $M$ and every $w \in NORM_M$, if $|A|_{M,w} = 1$ for all $A$ in $\Gamma$ then $|B|_{M,w} = 1$.

(Here what counts as a worlds model depends on which structural conditions (e.g. Weak Centering) have been imposed, so (VAL) really gives a family of notions of validity. Again, the restriction to normal worlds only makes a difference when one imposes structural requirements on the normal worlds of models that don't apply to all worlds.)[12]

We define $\vee$ from $\wedge$ and $\neg$, and $\exists$ from $\forall$ and $\neg$, and $\Diamond$ from $\Box$ and $\neg$, in the usual ways. ($|\Diamond A|_w$ is thus $\max\{|A|_x : x \in W_w\}$.)[13] $A \lhd\rhd B$ will abbreviate $(A \rhd B) \wedge (B \rhd A)$.

But there might be a reason to treat '$\rhd$' slightly differently. Many people, myself included, find it natural to suppose that $\neg(A \rhd B)$ should be to equivalent to $A \rhd \neg B$, modulo

12 We can generalize to the case where $B$ and the members of $\Gamma$ can contain free variables: for any model $M$, if $f$ is any function assigning $\{L^+\}_M$-names to free variables, and $A$ is any $L$-formula, let $A^f$ be the $\{L^+\}_M$-sentence that results from substitution by $f$. Then the generalization of (VAL) is

**(VAL$_{gen}$):** An inference from a set $\Gamma$ of $L$-formulas to an $L$-formula $B$ is *Burgess-valid* if for every worlds model $M$ and every $f$ for $M$ and every $w \in NORM_M$, if $|A^f|_{M,w} = 1$ for all $A$ in $\Gamma$ then $|B^f|_{M,w} = 1$.

13 Why take '$\Box$' as primitive, since $\Box A$ is equivalent to $\neg A \rhd A$? The answer is that the equivalence will be lost once we move to a 3-valued semantics, either because of the move to the modified Burgess evaluation procedure to be given next or to handle predicates like 'True'.

the assumption $\Diamond A$ (that is, each should imply the other on that assumption). We don't have that on the above semantics, unless we add strong assumptions (viz.: No Ties, No Incomparabilities and the Limit Condition); that was one of Stalnaker's arguments for imposing those assumptions. If we want that equivalence without the strong assumptions, we can get it by strengthening the 0 clause for '$\rhd$' while leaving the 1 clause as is. We then need a 3-valued framework to handle sentences that receive neither value 1 nor value 0. Our worlds models are still 2-valued for the moment, i.e. atomic sentences of $L^+$ can only take values in $\{0,1\}$, but we allow an additional value $\frac{1}{2}$ for conditionals and sentences containing them as components. The evaluation clause for '$\rhd$' is as follows:

**Modified Burgess evaluation procedure** [14]

$$|A \rhd B|_w = \begin{cases} 1 & \text{iff } (\forall x \in W_w)[|A|_x = 1 \supset (\exists y \leq_w x)[|A|_y = 1 \wedge \\ & (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 1)]] \\ 0 & \text{iff } (\forall x \in W_w)[|A|_x = 1 \supset (\exists y \leq_w x)[|A|_y = 1 \wedge \\ & (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 0)] \wedge (\exists x \in W_w)(|A|_x = 1) \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

(The 0-clause says that there are $A$-consistent $w$-neighborhoods, and each such has $A$-consistent sub-$w$-neighborhoods throughout which if $|A|$ is 1 then $|B|$ is 0.) I've already written the evaluation clauses for $\neg$, $\wedge$, $\forall$ and $\Box$ in a way that carries over automatically to allow for the extra value. (These clauses are called the *Strong Kleene rules*.)

The crucial thing about this alternative evaluation procedure for $\rhd$ is that if $|\Diamond A|_w$ is 1, i.e. if $(\exists x \in W_w)(|A|_x = 1)$, then $|\neg(A \rhd B)|_w$ is just $|A \rhd \neg B|_w$. Of course, a consequence will be a minimal non-classicality: excluded middle can fail for sentences containing '$\rhd$'. The cost of this isn't that high, I think: indeed, once we introduce a truth predicate, we'll need excluded middle to fail even more broadly than that.

What notion of validity goes with this modified evaluation scheme? There are several possible choices, but the one I will work with carries over the wording of (VAL) (or more generally, the (VAL$_{gen}$) of note 12) to the 3-valued case: validity involves preservation of value 1 at all normal worlds in all models (with the values now given by the modified evaluation rules).

In adding 'True' to the language we will need to adapt either the original Burgess semantics or the modified Burgess semantics to 3-valued worlds models. A *3-valued worlds model* is just like a 2-valued one except that in clause (iv) we replace $\{0, 1\}$ with $\{0, \frac{1}{2}, 1\}$, so that atomic sentences as well as conditionals can receive value $\frac{1}{2}$.[15] So 2-valued models are a special case of 3-valued. The most straightforward adaptation to the presence of 'True' would be to simply use the Burgess or modified Burgess rules as written. But instead of doing precisely that I will proceed in a more roundabout way, which nonetheless is modeled on these rules and agrees with them entirely for conditionals whose antecedent and consequent don't contain 'True'.

---

14 There's no danger of this requiring that the same conditional get both value 0 and 1 at a world. For assume as an induction hypothesis that $A$ and $B$ each have a unique value at each world. (Actually we'll only need that $B$ does.) If $A \rhd B$ gets value 0 at $w$, then there must be a $y \in W_w$ for which $|A|_y$ is 1 and $(\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 0)$; and if it gets 1 there must be a $y* \leq_w y$ such that $|A|_{y*} = 1 \wedge (\forall z \leq_w y*)(|A|_z = 1 \supset |B|_z = 1)$. But these require that $|B|_{y*}$ is both 0 and 1, contrary to the induction hypothesis.

15 For present purposes I keep the earlier restriction on the assignment $=_w$ to '=', though with the third value it could be liberalized somewhat to allow for indeterminate identity.

Validity will be defined as before: preservation of value 1 at all normal worlds of all models (that meet whatever structural conditions such as Weak Centering that one has imposed). However, when $L$ contains a truth predicate we'll restrict the models used in the definition, to "arithmetically standard" models that treat the predicate 'True' in a certain way. The details are in Section 3.

**§3. Truth and satisfaction: the strategy.** Suppose that $L$ contains a truth predicate (more specifically, a predicate of truth in $L$).[16] To be of interest, $L$ will also need to have the resources to talk of the bearers of truth, i.e. sentences, and their syntactic properties. Or instead of syntactic objects, $L$ could just contain arithmetic; we could talk of truth relative to a Gödel numbering. A language with a satisfaction predicate (from which truth can be defined, but not in general conversely) is more interesting; but to have a useful satisfaction predicate we need to be able talk of finite sequences of arbitrary objects from the universe of discourse, which requires additional mathematical resources. Moreover, dealing with satisfaction involves some notational complexity that can be confusing. So to keep things simple I'll take $L$ to involve a truth predicate but not a satisfaction predicate. It is routine to generalize what follows from truth to satisfaction (when the extra mathematical resources are available in $L$).

Rather than building syntactic notions into $L$, I'll follow the Gödel numbering route: $L$ will contain the predicates 'natural number', 'is zero', 'is the successor of', 'is the product of', and '='. (I'll fix a Gödel numbering $g$ of $L$.) I'll also be concerned only with worlds models $M$ whose arithmetic part is standard (an $\omega$-model) and the same from world to world. That is, I'll assume that in every model and every world in it, $U_w$ is a superset of the set $N$ of natural numbers, and 'natural number' is assigned $N$ as its extension, and the other arithmetic vocabulary is interpreted in the standard way. I'll call worlds models meeting these restrictions *arithmetically standard*. It's natural to restrict to them since without some such restriction the Gödel numbering results in "non-standard syntactic expressions" that have infinitely many distinct sub-expressions. If in defining validity we restrict to arithmetically standard worlds models, the result is *$\omega$-validity* (or *validity in $\omega$-logic*); it is this rather than regular validity that I will be primarily concerned with.

Kripke 1975, at least the part dealing with the Kleene construction, was concerned with the possibilities for *naive* truth (and satisfaction), though in languages not containing $\rhd$. Here I will extend his results to languages containing $\rhd$.

I informally defined "naive theory of truth" in my introductory remarks, but I should be more precise. Let a formula $Y$ be *a Tr-equivalent of* a formula $X$ if there are (possibly multiple) $L$-sentences $A$ such that $Y$ results from $X$ by (possibly multiple) substitutions of $True(\langle A\rangle)$ for $A$ and/or vice versa. A *naive theory of truth* is one where whenever $Y$ is a Tr-equivalent of $X$, $Y$ follows from $X$ and vice versa (i.e. the inferences from $X$ to $Y$ and $Y$ to $X$ are valid). The semantic paradoxes show that naivety is unattainable in classical logic, but Kripke (in his Kleene-based construction) showed it attainable in non-classical logic, by the use of 3-valued models. (Again, his language didn't contain $\rhd$.)

Naivety is not the sole requirement we should impose on a theory of truth: we also want it to obey reasonable compositional laws, and to allow the truth predicate to appear in an induction rule. More on these shortly.

16 Not in $L^+$: the new names in $L^+$ aren't part of the language $L$ for which we're giving a truth theory, and are dependent on a particular model of $L$. Any apparent loss in restricting truth to $L$-sentences should be met by generalizing from truth to satisfaction, as discussed later in the paragraph.

Our theory of truth should of course also be consistent, at least Post-consistent: that is, it shouldn't imply everything. I don't *in principle* require negation-consistency, i.e. the restriction to theories that for no $A$ imply both $A$ and $\neg A$. However, as is implicit in my earlier definition of validity, the theories I'll be developing satisfy disjunctive syllogism ($A \vee B, \neg A \models B$), and for those theories Post-consistency requires negation-consistency. (While there are familiar "paraconsistent" logics that avoid paradoxes without restricting excluded middle, by restricting disjunctive syllogism instead, they don't seem to me a promising framework for my ultimate goal of restricted quantification: the comments in Section 7 below on Beall *et al* 2006 and Beall 2009 may be enough to give some sense of this.)

Actually we want our naive truth theory to be more than (Post- or negation-) consistent: a consistent theory might, after all, imply the defeat of the Paris Commune, and no logic of truth should do that. What we want is for our theory of truth to be "consistent with any arithmetically standard worlds model" of the 'True'-free fragment of $L$, which I'll call $L_0$. More fully,

**GOAL:** We want to generate from each 2-valued arithmetically standard worlds model $M_0$ for $L_0$ a corresponding 3-valued worlds model $M$ for $L$ that (a) validates naive truth and (b) is exactly like $M_0$ except that it assigns a 3-valued extension to 'True'. It follows from (b) that the sentences of $L_0^+$ get the same value at $w$ in $M$ as in $M_0$, for each world $w$; and also that $M$ is arithmetically standard, given that $M_0$ is.

I'll take the allowable worlds models $M$ of $L$ to be just the ones generated from worlds models $M_0$ of $L_0$ in this way; that is, validity, consistency etc. *in the logic of truth* are defined by quantification over the arithmetically standard worlds models $M_0$ of the 'True'-free fragment of the language, and extending the valuation to sentences with 'True' by a procedure to be given.[17] (It isn't immediately obvious what this procedure should be when it comes to sentences containing both '$\rhd$' and 'True': e.g. to take a very simple Curry-like case, it isn't immediately obvious how to evaluate a sentence $K_\rhd$ constructed by the usual Gödel-Tarski techniques so as to be equivalent to $True(\langle K_\rhd \rangle) \rhd \neg True(\langle K_\rhd \rangle)$. Indeed I will consider several alternative procedures for constructing the extension.)

Note that if we can establish (GOAL), we get a kind of conservativeness result: letting *-consistency be consistency in $\omega$-logic, we have that any classically *-consistent set of sentences of $L_0$ is *-consistent in a naive truth theory.[18] The naive truth theory in question includes not merely the inferences from any sentence to its Tr-equivalents, it can include any other law validated in the construction of $M$ from $M_0$. What laws these are will of course depend on the details of the construction of $M$ from $M_0$, which is yet to be given.

But whatever the details, it is clear in advance that if (GOAL) is achieved then our construction will not only be one on which truth is naive, but one where mathematical induction in the form $A(0) \wedge (\forall n \in N)(A(n) \supset A(n+1)) \models (\forall n \in N)A(n)$ is legitimate even for formulas containing 'True'.[19] The reason is that in any arithmetically standard

[17] A slightly more general procedure will be mentioned in note 30.

[18] Calling this a conservativeness result could be misleading: there is no *deductive* conservativeness, it is a kind of semantic conservativeness in $\omega$-logic. Its purpose, as I've said, is to ensure that the set of principles to be declared valid in the naive truth theory is not merely consistent, but consistent with any set of assumptions in the 'True'-free language that are compatible with the conditional logic and standard models of arithmetic.

[19] Analogous forms with other modus-ponens-obeying conditionals in place of the '$\supset$' are guaranteed too.

worlds model, when the premises of this induction rule hold at a world the conclusion must too, and the construction guarantees that the new worlds model is arithmetically standard.

It is almost as immediate that the construction will validate the desirable composition principles, e.g.

**COMPOS-GENERAL:** $\forall x \forall y \forall z$(If $x$ and $y$ are sentences and $z$ is the result of applying '$\rhd$' to $x$ and $y$ in that order, then $\Box[True(z)$ if and only if $(True(x) \rhd True(y)]$).

For as long as the logic validates each instance of "$\Box$ [$A$ if and only if $A$]", then the naivety of truth guarantees the validity of each instance of

**COMPOS-SCHEMA:** $\Box[True(\langle A \rhd B\rangle)$ if and only if $(True(\langle A\rangle) \rhd True(\langle B\rangle))]$;

and since the constructed model is arithmetically standard, the generalization is guaranteed to hold in the model when the instances do. [This holds on *any* reading of 'if and only if', as long as "$\Box$ [$A$ if and only if $A$]" is validated. At the moment, the only available reading is '$\lhd\rhd$', but I will later add other biconditionals, and the point applies equally to them.]

**§4. Truth and satisfaction: the details.** I now outline a generalization of Kripke's construction. The initial generalization, which takes '$\rhd$' as a black-box, is completely routine, hardly a generalization at all; but a non-Kripkean ingredient is then required, to give a substantial account of '$\rhd$'.

Let's get the pure Kripke part of the construction out of the way first. It's clear from what has already been said that each of the worlds $w$ in the model for $L$ will be evaluated in part on the basis of $U_w$ and the $w$-extensions of $L_0$-predicates. The additional ingredients needed to evaluate $L^+$-sentences at each $w$ are:

- a 3-valued extension $T_w$ for 'True': it assigns values in $\{0, \frac{1}{2}, 1\}$ to objects in $U$. (We'll want it to assign non-zero values only to those objects that are Gödel numbers of $L$-sentences under the chosen Gödel numbering.)
- a function $j_w$ that assigns to each $L^+$-sentence of form '$A \rhd B$' a value in $\{0, \frac{1}{2}, 1\}$.

Let $T$ and $j$ be the functions that assign to each $w \in W$ a $T_w$ and $j_w$. Relative to any such $T$ and $j$, the Kleene rules tell us how to evaluate every $L^+$-sentence at $w$:

- For $p$ other than 'True', $|p(c_1, ..., c_k)|_{w,j,T}$ is just $p_w(o_1, ...o_k)$;
- $|True(c)|_{w,j,T}$ is $T_w(o)$, where $o$ is the object denoted by the $L^+$-name $c$;
- $|\neg A|_{w,j,T}$ is $1 - |A|_{w,j,T}$
- $|A \wedge B|_{w,j,T}$ is min $\{|A|_{w,j,T}\ , |B|_{w,j,T}\ \}$
- $|\forall x A|_{w,j,T}$ is min$\{|A(c/x)|_{w,j,T}$ : all $c$ that name members of $U_w\}$
- $|\Box A|_{w,j,T}$ is min$\{|A|_{w,j,T} : x \in W_w\}$
- $|A \rhd B|_{w,j,T} = j_w(A \rhd B)$.

The important thing about this is a monotonicity principle. Let $T \leq_K T^*$ mean that for every $w$ and every $L$-sentence $S$, if $T_w(S) = 1$ then $T^*_w(S) = 1$ and if $T_w(S) = 0$ then $T^*_w(S) = 0$. Then

**(MONOT):** For any $M$ and $j$: if $T \leq_K T^*$ then for any $w \in W$ and any $L^+$-sentence $A$, if $|A|_{w,j,T} = 1$ then $|A|_{w,j,T^*} = 1$ and if $|A|_{w,j,T} = 0$ then $|A|_{w,j,T^*} = 0$.

This is easily proved by an induction on the complexity of $A$. (The result is familiar from Kripke 1975, except that I've added a trivial $\rhd$ clause and a world-argument for $T$.)

This is the background for

PROPOSITION. ***[Kripke's observation.]*** *For any $M$ and $j$, there are $T$ ("Kripke fixed points" relative to $M$ and $j$) for which, for each $w \in W$:*

*For every $L$-sentence $A$, $|A|_{w,j,T} = T_w(g(A))$ [and hence $|A|_{w,j,T} = |True(c)|_{w,j,T}$, where $c$ denotes $g(A)$]; and*

*$T_w(o)$ is 0 if $o$ is not $g(A)$ for some $L$-sentence $A$.*

*In particular, for any $M$ and $j$ there is a minimal fixed point $T_{min}$, i.e. a fixed point (relative to $M$ and $j$) such that for every other fixed point $T$ (relative to $M$ and $j$), $T_{min} \leq_K T$.*

Kripke's observation is easily proved by transfinite induction.[20]

It easily follows that as long as $j$ is *transparent*, in the sense that it assigns Tr-equivalent formulas the same value, then the naivety condition is met: whenever $A$ and $B$ are Tr-equivalent, $|A|_{w,j,T_{min}} = |B|_{w,j,T_{min}}$. (And similarly for fixed points $T$ other than $T_{min}$.)

The definition of $T_{min}$ depended on the choice of $M$ and $j$, but given those, $T_{min}$ is uniquely determined; so we can abbreviate $|A|_{w,j,T_{min}}$ as $|A|_{w,j}$. To repeat, this valuation yields naive truth as long as $j$ is transparent.

The harder task is to construct an appropriate transparent $j$-function for evaluating conditionals at worlds. What we want is a transparent $j$ that leads to a logic that reduces to the Burgess or modified-Burgess logic when applied to 'True'-free sentences *and which weakens the laws as little as possible when sentences with 'True' are allowed as instances.* There are at least two approaches to constructing such a $j$ function: a revision construction, with similarities to those in Field 2008; or a fixed point construction, with similarities to those in Field 2014.

The revision construction is simpler, so I'll focus on it, but will also make a few remarks about the (perhaps more aesthetically pleasing) fixed point construction.

***4.1. The revision construction.*** Fix a worlds model $M_0$ for $L_0$. Suppose we have given a provisional valuation $j_v$, which assigns values $|B \rhd C|_{w,j_v}$ to any $L^+$-sentences $B$ and $C$. As we've seen, this indirectly gives a value $|A|_{w,j_v}$ to every $L^+$-sentence $A$ at every world, via the Kripke minimal fixed point construction; let's just write this as $|A|_{w,v}$. We want to use this valuation $j_v$ to construct a revised one $j_{v+1}$, perhaps a better one, which is transparent if the original one is; the structure of worlds is used in the revision.

There are two possibilities for $j_{v+1}$, one based on the original Burgess valuation rules and the other based on the variant. For the original it is:

$$j_{w,v+1}(A \rhd B) \text{ is } \begin{cases} 1 & \text{iff } (\forall x \in W_w)[|A|_{x,v} = 1 \supset (\exists y \leq_w x)[|A|_{y,v} = 1 \wedge \\ & (\forall z \leq_w y)(|A|_{z,v} = 1 \supset |B|_{z,v} = 1)]] \\ 0 & \text{iff } (\exists x \in W_w)[|A|_{x,v} = 1 \wedge (\forall y \leq_w x)[|A|_{y,v} = 1 \supset \\ & (\exists z \leq_w y)(|A|_{z,v} = 1 \wedge |B|_{z,v} = 0)]] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

[20] Holding $M$ and $j$ fixed, we define $T_0$ to be the function assigning the value $\frac{1}{2}$ to every Gödel number of an $L$-sentence, and 0 to everything else; $T_{\sigma+1}$ the function assigning every world $w$ and $L$-sentence $A$ the value $|A|_{w,j,T_\sigma}$; and $T_\lambda$ (for limit $\lambda$) the function assigning every world $w$ and $L$-sentence $A$ the value

$$\begin{cases} 1 & \text{if for some } \sigma<\lambda \text{ and every } \tau \text{ such that } \sigma \leq \tau<\lambda,\ |A|_{w,j,T_\tau} = 1; \\ 0 & \text{if for some } \sigma<\lambda \text{ and every } \tau \text{ such that } \sigma \leq \tau<\lambda,\ |A|_{w,j,T_\tau} = 0; \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

We can then easily prove by induction that if $\sigma < \tau$, $T_\sigma \leq_K T_\tau$. Cardinality considerations then show that there are ordinals $\sigma$ (of the cardinality of $U_M$) after which the assigned $T$ never changes. Taking $T_{min}$ to be $T_\sigma$ for such a $\sigma$, we get the desired result.

For the variant, it's the same except for a modified 0 clause:

$$0 \text{ iff } (\forall x \in W_w)[|A|_{x,\nu} = 1 \supset (\exists y \leq_w x)[|A|_{y,\nu} = 1 \wedge (\forall z \leq_w y)(|A|_{z,\nu} = 1 \supset |B|_{z,\nu} = 0)]] \wedge (\exists x \in W_w)(|A|_{x,\nu} = 1).$$

Choose whichever you like: the construction that follows works with either choice.

To get the revision process started, we need a starting valuation $j_0$, and we want it to be transparent since this will guarantee that later $j_\nu$ are as well. For simplicity I'll take a trivial $j_0$, which assigns value $\frac{1}{2}$ to each conditional at each world. It makes little difference, because the effect of the starting values gets *almost* completely wiped out as the construction proceeds. (It gets *completely* wiped out for sentences not containing 'True': whatever the starting values, any such sentence gets the value that it gets in the 2-valued worlds model for the corresponding version of Burgess semantics by stage $n$, where $n$ is the maximum depth to which '$\rhd$' is embedded in the scope of other '$\rhd$'s in $A$; and it keeps that value at all subsequent stages. So from stage $\omega$ on, all 'True'-free sentences get "the value they should", whatever the starting valuation.)

Finally, we need a policy on limit stages. Here the choice is important, and we choose continuity with respect to 1 and 0. That is, if $\lambda$ is a limit ordinal then for any world $w$ and any conditional $A \rhd B$, $j_\lambda$ assigns the conditional 1 at a world if and only if for some $\mu < \lambda$, for every ordinal $\nu$ in the open interval $(\mu, \lambda)$ assigns the conditional value 1 at that world; and similarly for 0. (So "irregularity arbitrarily close to $\lambda$" at a world as well as "constant $\frac{1}{2}$ sufficiently close to $\lambda$" at that world lead to value $\frac{1}{2}$ at $\lambda$ at that world.)

We can summarize these choices in a single definition. For the semantics based on the modified Burgess, which I prefer, it's

$$j_{w,\kappa}(A \rhd B) \text{ is } \begin{cases} 1 & \text{if } (\exists \mu < \kappa)(\forall \nu \in [\mu, \kappa))(\forall x \in W_w)[|A|_{x,\nu} = 1 \supset (\exists y \leq_w x) \\ & [|A|_{y,\nu} = 1 \wedge (\forall z \leq_w y)(|A|_{z,\nu} = 1 \supset |B|_{z,\nu} = 1)]] \\ 0 & \text{if } (\exists \mu < \kappa)(\forall \nu \in [\mu, \kappa))[(\forall x \in W_w)[|A|_{x,\nu} = 1 \supset (\exists y \leq_w x) \\ & [|A|_{y,\nu} = 1 \wedge (\forall z \leq_w y)(|A|_{z,\nu} = 1 \supset |B|_{z,\nu} = 0)]] \wedge \\ & (\exists x \in W_w)(|A|_{x,\nu} = 1)] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

($[\mu, \kappa)$ is the half-open interval of $\nu$ such that $\mu \leq \nu < \kappa$.) For the semantics based on the original Burgess, modify the 0 clause in the obvious way.

It's evident that on either variant, each $j_\kappa$ is transparent if all preceding $j_\nu$ are transparent; so by transfinite induction, all are transparent.

At each world, all 'True'-free sentences get the desired value (i.e. the one given in the 2-valued model from which we started) by stage $\omega$, and keep it at later stages. But there is much greater irregularity for sentences containing 'True', due to the interaction between 'True' and '$\rhd$'.[21] In particular there is no fixed point. How then are we to select a privileged $j$?

The sequence of $j_\nu$ is a revision sequence in the sense of Gupta and Belnap 1993. (The revision sequence depends on the model $M_0$, as well as on the choice of Burgess or modified-Burgess.) One well-known feature of revision sequences is that there are evaluations $j$ that appear arbitrarily late in the revision process; indeed, there are ordinals $\kappa$ such that for any $\mu \geq \kappa$ and any $\zeta$, there is a $\nu \geq \zeta$ such that $j_\nu = j_\mu$.[22] Call any

[21] The sentence itself needn't even contain '$\rhd$' for the irregularity to occur, because the use of 'True' typically makes other sentences relevant to the evaluation.

infinite such $\kappa$ *final* (relative to model $M_0$),[23] and let $FIN$ (or $FIN_{M_0}$) be the class of final ordinals.

But not all final ordinals assign the same $j$ (if they did, it would be a fixed point). Which to pick? Obviously we want one that will yield as nice laws for $\rhd$ as possible. Gupta and Belnap 1993 have a general theorem, their Reflection Theorem, that we can bring to bear. Applied to this case, that theorem says:

PROPOSITION. ***[Gupta-Belnap]*** *There are limit ordinals $\Omega$ ("reflection ordinals for the sequence $j_\kappa$")*[24] *such that*

(i) *$\Omega$ is final*

(ii) *For any $L^+$-formulas A and B, and any world $w$ and any $d \in \{0, \frac{1}{2}, 1\}$,*

*$(\exists\mu < \Omega)(\forall\nu \in [\mu, \Omega))((j_{w,\nu}(A \rhd B) = d)$ if and only if $(\forall\nu \in FIN)(j_{w,\nu}(A \rhd B) = d)$.*

Moreover, in the above semantics these reflection ordinals have an especially useful property:

PROPOSITION. ***[Fundamental Theorem for L (revision-theoretic version).]*** *For any reflection ordinal $\Omega$, any $w \in W$, and any $L^+$-sentence A,*

*(a) $|A|_{w,\Omega} = 1$ if and only if $(\forall\nu \in FIN)(|A|_{w,\nu} = 1)$*

*and (b) $|A|_{w,\Omega} = 0$ if and only if $(\forall\nu \in FIN)(|A|_{w,\nu} = 0)$.*

Since there is only one possible value other than 0 and 1, these two clauses imply that each reflection ordinal $\Omega$ is associated with the same $j_\Omega$. This $j_\Omega$ is the valuation for $\rhd$-conditionals that I'll be employing, e.g. in determining validity.

The Fundamental Theorem as stated here is similar to that given in Field 2008, but the conditional there was different. The proof given there included a proof of [Gupta-Belnap], since I was unaware of their theorem at the time. (Belated apologies to them for not being able to give credit.) A proof of the Fundamental Theorem for the language of this paper, now relying on [Gupta-Belnap] to save work, is given in Appendix A.

Note that when $A$ is a conditional $B \rhd C$, the 1-clause of the Fundamental Theorem together with the evaluation rules for $\rhd$ yield that for any reflection ordinal $\Omega$ and $w \in W$,

**1-clause:** $|B \rhd C|_{w,\Omega} = 1$ if and only if $(\forall\nu \in FIN)(\forall x \in W_w)[|B|_{x,\nu} = 1 \supset (\exists y \leq_w x)[|B|_{y,\nu} = 1 \wedge (\forall z \leq_w y)(|B|_{z,\nu} = 1 \supset |C|_{z,\nu} = 1)]]$.

Since $\Omega \in FIN$, this yields a necessary but not sufficient condition for $|B \rhd C|_{w,\Omega} = 1$ that involves no ordinals other than $\Omega$:

**1-clause Corollary:** If $|B \rhd C|_{w,\Omega} = 1$ then $(\forall x \in W_w)[|B|_{x,\Omega} = 1 \supset (\exists y \leq_w x)[|B|_{y,\Omega} = 1 \wedge (\forall z \leq_w y)(|B|_{z,\Omega} = 1 \supset |C||_{z,\Omega} = 1)]]$.

That is, since we've chosen to use $j_\Omega$ for our final valuation: the 1-clause we've adopted is strictly stronger than the 1-clause of the Burgess and modified-Burgess semantics.

22 Since the revision sequence here is *Markovian* in the sense that for any ordinals $\mu$, $\kappa$ and $\nu$, if $j_\mu = j_\kappa$ then $j_{\mu+\nu} = j_{\kappa+\nu}$, we can simplify to: for any $\zeta$, there is a $\nu \geq \zeta$ such that $j_\nu = j_\kappa$. If this holds for $\kappa$ in a Markovian sequence, it is bound to hold for any $\mu > \kappa$.

23 It isn't really necessary to demand infinitude explicitly, it's entailed by the rest, as the reader can easily prove using 'True'-free sentences where '$\rhd$' is embedded to depth $n$ for arbitrarily large $n$.

24 Which ordinals are reflection ordinals will depend on the starting model $M_0$.

ZU064-05-FPR S1755020315000301 19 September 2015 12:42



But since all final ordinals are infinite, all 'True'-free sentences receive the same value at all final ordinals; this means that for such $B$ and $C$ the 'if...then' in the corollary becomes an 'if and only if'. In other words, we're guaranteed that *the Burgess/modified-Burgess 1-clause is retained for 'True'-free sentences.*

Moreover, as long as we have Weak Centering at $w$, the 1-clause corollary yields the following for all $B$ and $C$ (not just the 'True'-free ones):

**Modus Ponens for $\triangleright$:** If $|B \triangleright C|_{w,\Omega} = 1$ and $|B|_{w,\Omega} = 1$ then $|C|_{w,\Omega} = 1$.

(The label 'Modus Ponens' is really appropriate only if we have Weak Centering at all normal $w$.)

Something similar holds for the 0-clause, though the details depend on which version of the 0 clause one uses. In both cases, we get strictly stronger conditions than would be given by direct application of the Burgess or modified Burgess rules: e.g. for the semantics based on modified Burgess we get

$$\text{If } |B \triangleright C|_{w,\Omega} = 0 \text{ then } (\forall x \in W_w)[|B|_{x,\Omega} = 1 \supset (\exists y \leq_w x)[|B|_{y,\Omega} = 1 \wedge (\forall z \leq_w y)(|B|_{z,\Omega} = 1 \supset |C|_{z,\Omega} = 0)]] \wedge (\exists x \in W_w)(|B|_{x,\Omega} = 1).$$

But again, when confined to 'True'-free sentences the 'if' becomes an 'if and only if': *the Burgess or modified Burgess 0 clause is also retained for 'True'-free sentences.*

(When $w$ is weakly centered, the above yields

**0 Law for $\triangleright$:** If $|B \triangleright C|_{w,\Omega} = 0$ and $|B|_{w,\Omega} = 1$ then $|C|_{w,\Omega} = 0$ (and indeed, $|C|_{x,\Omega} = 0$ whenever $x \sim_w w$),

which also strikes me as desirable but will play no role in what follows. Had we based the semantics on the original Burgess, we'd have needed that $w$ be strongly centered to get this result.)

***4.2. Where are we?*** For each starting arithmetically standard worlds model $M_0$ for the 'True'-free fragment $L_0$ of $L$ (with $\triangleright$ evaluated either by the standard Burgess or variant Burgess rules), we have chosen a transparent $j_\Omega$ to evaluate all $L^+$-conditionals at each world (including those containing embedded conditionals and/or 'True'), and a $T$ to evaluate truth-claims at each world. The worlds, and their division into normal and non-normal, are the same in the new model as in the old. (In particular, if the old contains no non-normal worlds, the new one won't either.) The assignment of accessibility sets $W_w$ and pre-orders $\leq_w$ is also the same in the new model as in the old; so are the assignments of extensions to predicates at each world. And at each world, $j_\Omega$ assigns the same values to 'True'-free conditionals (and hence 'True'-free sentences more generally) as the original model on $M_0$ did. Finally, by the transparency of $j$ and the features of the Kripke construction, the truth predicate is naive; and since the model is arithmetically standard, there can be no worry about using formulas with 'True' in the induction rule or validating generalities (e.g. composition rules) whose instances are valid.[25]

[25] A feature of the model as described is that it is not value-functional: the value of $A \triangleright B$ at a world isn't determined wholly by the values of $A$ and $B$ at it and other worlds. The reason is that all these values are values at a reflection $\Omega$, and these depend on values at all non-reflection ordinals in $FIN$. But it isn't hard to use what's been done here to construct an enriched value space (along the lines of Field 2008, Section 17.1) in which we do have value-functionality: the value space for that will have infinitely many values, not linearly ordered. (The space is a set of functions from an initial segment of the ordinals to $\{0, \frac{1}{2}, 1\}$, where the length of the initial segment is the

The following are laws of this construction: by which I mean, schemas all of whose instances are valid (whatever structural conditions, such as Weak Centering at normal worlds, we decide on):

- $A \rhd A$
- $[A \rhd (B \wedge C)] \lhd\rhd [(A \rhd B) \wedge (A \rhd C)]$
- $[(A \rhd C) \wedge (B \rhd C)] \rhd [(A \vee B) \rhd C]$
- $[A \rhd (B \wedge C)] \rhd [(A \wedge B) \rhd C]$.

These are laws both when the evaluation rule for $\rhd$ is based on the original Burgess rule and when it is based on the modified rule: the 0 clause makes no difference. Indeed on both constructions they are all *strong laws*, by which I mean that their instances have value 1 at all worlds of every model, not just all normal worlds. That's important because it means that the result of prefixing any string of $\Box$s and $\Diamond$s to one of these is also a law. Related, it guarantees other "regular behavior", such as that we can strengthen antecedents in the laws. That is, even though we don't want and don't get that $Y \rhd Z$ entails $X \wedge Y \rhd Z$ for variably strict conditionals, still if $Y \rhd Z$ is a strong law then so is $X \wedge Y \rhd Z$ (even if $X$ is true only at non-normal worlds). Similarly, if $X \rhd Y$ and $Y \rhd Z$ are strong laws then so is $X \rhd Z$.[26] Proving that the bulleted schemas are strong laws is straightforward.[27] Note that since $\Box(A \lhd\rhd A)$ is valid, then by naivety so are $\Box(True(\langle A \rangle) \lhd\rhd A)$ and $\Box(\neg True(\langle A \rangle) \lhd\rhd \neg A)$, and hence

distance between successive reflection ordinals.) But for purposes of this paper there's no need for value-functionality.

[26] The proof that "antecedent strengthening" and transitivity are legitimate for strong laws uses the Fundamental Theorem as applied to $\rhd$-sentences. Let $W^*$ be the set of worlds that are $n$-accessible from worlds for some $n$. (On reasonable assumptions this will just be $W$, but the proof doesn't need this.) For antecedent strengthening, suppose that $Y \rhd Z$ has value 1 at all worlds at reflection ordinals. Then it has value 1 at all worlds at all final ordinals, which means that at all final ordinals and all worlds in $W^*$, if $Y$ has value 1 then so does $Z$; and that includes all worlds where $X$ has value 1. From this it's evident that $X \wedge Y \rhd Z$ has value 1 at all worlds in $W$ (even those not in $W^*$, since only those in $W^*$ are accessible to them) at all final ordinals, and in particular at reflection ordinals. The argument for transitivity is similar.

[27] The key observation for all of them is that for $|X \rhd Y|_{w,\Omega}$ to be 1, it suffices that for all worlds $w^*$ and all final ordinals $\nu$, if $|X|_{w^*,\nu} = 1$ then $|Y|_{w^*,\nu} = 1$. Given that, it's simply a matter of relativizing the proof that one would give for the Burgess-based semantics in the ground level language to a given $\nu$. For instance, for the right to left direction of the second listed law: Suppose that $|A \rhd B|_{w^*,\nu} = |A \rhd C|_{w^*,\nu} = 1$. Then for every $x$ in $W_{w^*}$ such that $|A|_{w^*,\nu} = 1$, there is a $y_1 \leq_{w^*} x$ such that

$$\text{(a) } |A|_{y_1,\nu} = 1 \wedge (\forall z \leq_{w^*} y_1)[|A|_{z,\nu} = 1 \supset |B|_{z,\nu} = 1],$$

and for every $y_1$ in $W_{w^*}$ such that $|A|_{y_1} = 1$, there is a $y_2 \leq_{w^*} y_1$ such that

$$\text{(b) } |A|_{y_2,\nu} = 1 \wedge (\forall z \leq_{w^*} y_2)[|A|_{z,\nu} = 1 \supset |C|_{z,\nu} = 1].$$

Since $\leq_{w^*}$ is a pre-order on $W_{w^*}$, (a) entails its analog (a*) where $y_2$ replaces $y_1$; and that with (b) yields

$$|A|_{y_2,\nu} = 1 \wedge (\forall z \leq_{w^*} y_2)[|A|_{z,\nu} = 1 \supset |B \wedge C|_{z,\nu} = 1],$$

which entails $|A \rhd B \wedge C||_{w^*,\nu} = 1$.

(This proof and the proofs of the other laws just given doesn't depend on the use of a reflection ordinal for our evaluation: that should be no surprise, since the Fundamental Theorem shows that a single sentence can only have value 1 at reflection ordinals if it has value 1 at all final ordinals. Where the fact that validity requires preservation of value 1 only at reflection ordinals is important is for inferences from premises: e.g. Modus Ponens (assuming Weak Centering at normal worlds) and $A \wedge B \vDash A \rhd B$ (assuming Strong).)

$\Box(True(\langle\neg A\rangle) \lhd\rhd \neg True(\langle A\rangle))$. (And by the remarks at the end of Section 3, this means that we have a general composition principle for negation: for any sentence $x$, the negation of $x$ is true if and only if $x$ is not true.)

The fact that the above laws all hold in the construction with naive truth is interesting, because these are exactly the axiom schemas that Burgess uses in the quantifier-free case for the 'True'-free fragment of the language. He gives a completeness proof there, for a system with these axioms, a necessitation rule, and the rule that for any string P of $\Box$s and $\Diamond$s, if $\models P\Box(A \equiv B)$ then $\models P[(A \rhd C) \equiv (B \rhd C)]$. The last rule is inappropriately weak in the 3-valued framework: we want a rule that has bite even when $A$ and $B$ aren't bivalent. (An adequate replacement requires the additional conditional '$\rightarrow$' soon to be introduced).[28] More generally, because the 3-valued background is weaker, the Burgess axiomatization doesn't give a complete proof-procedure in the 3-valued context.[29] Still, I think that the fact that his axioms carry over unchanged is some indication that adding a naive truth predicate hasn't seriously compromised the laws of '$\rhd$' (and once we add the '$\rightarrow$' things will look even better).

In addition, we've seen that as long as we restrict the ground models to those with Weak Centering at normal worlds (as is required for Modus Ponens in the ground language), then Modus Ponens for $\rhd$ also holds in the expanded logic with 'True'. (Some of the laws obtained in the 2-valued logic by adding restrictions on the $\leq_w$ can only be carried over straightforwardly to the full logic with 'True' when stated using the aforementioned conditional '$\rightarrow$' that generalizes the material conditional. We'll turn to that conditional in Section 5.)

That's the revision construction.

***4.3. The fixed point construction.*** As I've mentioned, one can also give a fixed point construction that yields a rather similar outcome. Again consider valuation functions $j$ that assign values in $\{0, \frac{1}{2}, 1\}$ to each pair of a world and a $\rhd$-conditional; again we're only interested in valuation functions that are transparent. The idea is to show that there is a set $\mathbf{J}$ of transparent valuations, with a distinguished member $j^*$, where we have

PROPOSITION. *[**Fundamental Theorem for** $L$ **(fixed point version).**] For any $w \in W$, and any $L^+$-sentence $A$,*

(a) $|A|_{w,j^*} = 1$ *if and only if* $(\forall h \in \boldsymbol{J})(|A|_{w,h} = 1)$

(b) $|A|_{w,j^*} = 0$ *if and only if* $(\forall h \in \boldsymbol{J})(|A|_{w,h} = 0)$.

So $j^*$ plays more or less the role that the $j_\Omega$ for reflection $\Omega$ play in the revision approach, and $\mathbf{J}$ plays more or less the role of the set of those $j$ that occur arbitrarily late in the revision process (i.e. at ordinals in $FIN$). Here too, the various valuations $j$ get a semantics whereby for any $L^+$-sentences $A$ and $B$ and any world $w$, $j(w, A \rhd B)$ is determined in a natural way from the values that valuations related to $j$ give to $B$ in worlds near $w$ where $A$ has value 1; and the semantics gives the values in the original model to $L^+$-sentences not

---

[28] The best replacement is:

$$\text{For any string } P \text{ of } \Box\text{s and } \Diamond\text{s, if} \models P\Box(A \leftrightarrow B) \text{ then} \models P[(A \rhd C) \Leftrightarrow (B \rhd C)];$$

here '$\Rightarrow$' is defined from '$\rightarrow$' and strengthens it in a way to be discussed in Section 5, and '$\leftrightarrow$' and '$\Leftrightarrow$' are defined from '$\rightarrow$' and '$\Rightarrow$' in the obvious ways. (The displayed law with mixed biconditionals entails the versions with two '$\leftrightarrow$' and with two '$\Leftrightarrow$'.)

[29] Indeed, the fact that we've restricted to arithmetically standard models immediately rules out the possibility of a complete proof procedure.

containing 'True'. To get the proper intersubstitutivity of logical equivalents, one needs to set up the semantics in a slightly non-obvious way. I sketch the construction in Appendix B; it is a generalization to variably strict conditionals of the one in Field 2014, and that paper will enable the reader to easily fill out the sketch in the Appendix.

(The basic idea of using a fixed point on a set of valuations was suggested in Yablo 2003; but Yablo's procedure didn't cut down the set of valuations quantified over in the semantics of each world nearly far enough—indeed, highly irregular valuations were included—and this led to extreme failure of intersubstitutivity of logical equivalents in embedded conditionals. Introducing chains in the manner of Appendix B seems to be the simplest acceptable way of accommodating Yablo's basic insight.)[30]

The remarks in Section 4.2 about the revision construction carry over to the fixed point construction virtually unchanged. In particular, the laws listed there are valid here too (again, with Modus Ponens as long as the original model has Weak Centering at normal worlds).

**§5. "Material-like" conditionals.** Many uses of 'if ... then' in English are captured reasonably well by a variably strict conditional like '$\rhd$', but some uses are more in line with a material conditional: in particular, the conditional used to restrict universal quantification is. "All $A$ are $B$" can't be rendered as $\forall x(Ax \rhd Bx)$: that's too strong when '$\rhd$' is an ordinary indicative (or subjunctive) conditional. For instance, "Everyone who will be elected President in 2016 is female" might be true but "For everyone $x$, if $x$ is elected President in 2016 then $x$ is female" presumably isn't: on the ordinary indicative reading, Jeb Bush and many others are counterexamples even if unelected. In a 2-valued context, we can represent "All $A$ are $B$" as $\forall x(Ax \supset Bx)$, where this is short for $\forall x(\neg Ax \vee Bx)$. But in a 3-valued context with restrictions on excluded middle, we can't use a $\supset$ defined in terms of $\neg$ and $\vee$ (at least if we want such schemas as "All $A$ are $A$" and "All $A$ are either $A$ or $B$" to be logical laws); we need a new conditional '$\rightarrow$' or '$\Rightarrow$', that reduces to $\supset$ for 2-valued sentences just as our '$\rhd$' reduces to the "classical" variably strict conditional.[31] I find it plausible that this quantifier-restricting conditional is contraposable, but I needn't insist on this: I will simply take '$\Rightarrow$' to be a contraposable conditional and '$\rightarrow$' to be a non-contraposable one, and we can leave open for now which of the two is to be used to define restricted quantification. There is no need for separate theories of '$\rightarrow$' and '$\Rightarrow$': we can take the basic conditional to be the non-contraposable '$\rightarrow$', and define $A \Rightarrow B$ as $(A \rightarrow B) \wedge (\neg B \rightarrow \neg A)$, which ensures that '$\Rightarrow$' is contraposable. The basic '$\rightarrow$' and the derived '$\Rightarrow$' have uses other than for restricting quantification: as observed in note 28, they are also needed for some of the

[30] Yablo's paper also suggests the use of multiple Kripke fixed points for 'True' instead of the minimal ones; that idea can be employed with any of the constructions for '$\rhd$' in this section, both revision-theoretic and fixed point, and has what are arguably some advantages. For further discussion (in a revision-theoretic context with a different conditional), see Field 2008, Section 17.5. Again, it doesn't matter to the issues of this paper whether one makes these modifications.

[31] I should note that the notation used in this paper is almost the reverse of the notation in Field 2014. There, the material-like conditional used to restrict quantification (which was assumed contraposable) was symbolized as $\blacktriangleright$, and $\rhd$ was its non-contraposable generalization; whereas $\rightarrow$ was used to symbolize a conditional with very much the flavor of the $\rhd$ used here, though it wasn't based on a Stalnaker-Lewis-Pollock-Burgess multiple worlds semantics. Sorry for any confusion, but I think the new notation distinctly better.

An alternative to introducing a new conditional and defining universal restricted quantification in terms of it is to take a binary restricted quantifier $(\forall x \ni Ax)Bx$ as primitive. One can define '$\Rightarrow$' (though not '$\rightarrow$') from it, as well as the other way around.

laws of '$\rhd$' (and for these purposes, '$\rightarrow$' as well as '$\Rightarrow$' is required). But though I'll take '$\rightarrow$' as basic, '$\Rightarrow$' will be the primary focus, because at least in my own view, it is this contraposable one that is ordinarily used to restrict universal quantification.

There are several options in the literature for such a conditional '$\rightarrow$' (or a corresponding contraposable '$\Rightarrow$'). Some of these are broadly like the revision-theoretic and fixed point options for $\rhd$ given in Section 4; but a key difference is that the valuations at a single world look only at other values at that same world.

For the moment let's ignore the interaction between '$\rightarrow$' and '$\rhd$', and focus on a language $L^*$ just like $L$ except that it has '$\rightarrow$' instead of '$\rhd$'. A language with both '$\rightarrow$' and '$\rhd$' is far more interesting, and will be treated in Section 6. That is what we'll need for a proper logic for restricted quantification in naive truth theory, a matter I'll turn to in Section 7. But for the moment, I look at $L^*$, which has '$\rightarrow$' only.

$L^*$, like $L$, contains 'True'; if it didn't, and could be given a 2-valued semantics, we could just define $\rightarrow$ from $\neg$ and $\vee$ in the usual way. As before, the semantics for 'True' will be given by Kripkean constructions in which valuations $v$ (analogous to the previous $j$) for '$\rightarrow$' at each world are held fixed; the real work then consists in the specification of an appropriate valuation for '$\rightarrow$' at each world.

A revision-theoretic construction of such a valuation for '$\Rightarrow$' was given in Field 2008; instead of what I called the "Official Conditional", given in Ch. 16, I now prefer the "first variation" given in Section 17.5, which modifies the 0 clause.[32] And I want to adapt it to the non-contraposable '$\rightarrow$'. Since $L^*$ contains '$\Box$', we need to add a worlds parameter; but the semantics for '$\rightarrow$' is given world-by-world, unlike for '$\rhd$', and is thus considerably simpler. It goes like this:

$$|A \rightarrow B|_{w,\alpha} = \begin{cases} 1 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta,\alpha))[|A|_{w,\gamma} = 1 \supset |B|_{w,\gamma} = 1] \\ 0 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta,\alpha))[|A|_{w,\gamma} = 1 \wedge |B|_{w,\gamma} = 0] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

If we then define '$\Rightarrow$' from '$\rightarrow$' as above, we get something similar but with a strengthened 1-clause:

$$|A \Rightarrow B|_{w,\alpha} = \begin{cases} 1 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta,\alpha))[|A|_{w,\gamma} \leq |B|_{w,\gamma}] \\ 0 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta,\alpha))[|A|_{w,\gamma} = 1 \wedge |B|_{w,\gamma} = 0] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Like the earlier construction with '$\rhd$', this construction gives rise to a set of final ordinals that include reflection ordinals $\Delta$, and a Fundamental Theorem just like the previous:

PROPOSITION. *[**Fundamental Theorem for $L^*$(revision-theoretic version).**] For any reflection ordinal $\Delta$, any $w \in W$, and any $L^+$-sentence $A$,*

(a) $|A|_{w,\Delta} = 1$ *if and only if* $(\forall\gamma \in FIN)(|A|_{w,\gamma} = 1)$

(b) $|A|_{w,\Delta} = 0$ *if and only if* $(\forall\gamma \in FIN)(|A|_{w,\gamma} = 0)$.

It can be shown that if the 'True'-free fragment $L_0^*$ is 2-valued, $\rightarrow$ and $\Rightarrow$ are each equivalent to the material conditional $\supset$ on $L_0^*$. (If the 'True'-free fragment $L_0^*$ is 3-valued, as it would be if we were to add $\rhd$ to the language and used the modified-Burgess-based semantics, then $\Rightarrow$ behaves on it like the Lukasiewicz 3-valued conditional, and $\rightarrow$ like a less familiar one.)

---

[32] This switch yields a cleaner relation between $|A \Rightarrow B \wedge C|$ on the one hand and $|A \Rightarrow B|$ and $|A \Rightarrow C|$ on the other: see the end of this section. That in turn is important for restricted quantifier law 4a* in Section 7.

ZU064-05-FPR S1755020315000301 19 September 2015 12:42



As with $\rhd$, only the valuations at reflection ordinals are relevant to validity: an inference is valid iff in all starting models and all worlds $w$ in them and all reflection $\Delta$, if the premises have value 1 at $w$ and $\Delta$ then so does the conclusion.

Alternatively, we could adapt the fixed point semantics, to get a set $\mathbf{R}$ of valuations $u$ assigning values in $\{0, \frac{1}{2}, 1\}$ to each $\rightarrow$-conditional at each world, with privileged member $v^*$. Again, the semantics for non-privileged members of $\mathbf{R}$ is given by a somewhat complicated chain construction analogous to that in Appendix B, but again it very much simplifies for $v^*$: we get

PROPOSITION. ***[Fundamental Theorem for L* (fixed point version).]*** *For any $w \in W$, and any $L^{*+}$-sentence $A$,*

(a) $|A|_{w,v*} = 1$ *if and only if* $(\forall u \in \mathbf{R})(|A|_{w,u} = 1)$

(b) $|A|_{w,v*} = 0$ *if and only if* $(\forall u \in \mathbf{R})(|A|_{w,u} = 0)$.

Only the special $v^*$ is used in the definition of validity.[33]

I note two consequences of the Fundamental Theorems for $L^*$:

**Modus Ponens for $\rightarrow$ and $\Rightarrow$:** $A, A \rightarrow B \vDash B$ (and hence $A, A \Rightarrow B \vDash B$)

**Weak Equivalence of $\neg(A \rightarrow B)$ and $\neg(A \Rightarrow B)$ to $A \wedge \neg B$:** The inference from either $\neg(A \rightarrow B)$ or $\neg(A \Rightarrow B)$ to $A \wedge \neg B$ is valid, and so are the reverse inferences.

Why is the second one called "Weak" Equivalence? Two reasons: (a) While (in the revision version) $|\neg(A \rightarrow B)|_{w,\Delta}$ (or $|\neg(A \Rightarrow B)|_{w,\Delta}$) is 1 iff $|A \wedge \neg B|_{w,\Delta} = 1$, there is no analogous claim for 0. (b) Even for 1, the result holds only for reflection $\Delta$, not for all final ordinals. (Similarly in the fixed point case: the equivalence holds only at $v^*$, not at all valuations in $\mathbf{R}$.) A consequence of (b) is that $\neg(A \rightarrow B)$ won't in general be intersubstitutable with $A \wedge \neg B$ even in positive contexts, unless those contexts are outside the scope of $\rightarrow$'s.

The proofs of Modus Ponens and Weak Equivalence are routine applications of the Fundamental Theorem (for the appropriate construction) together with the evaluation clauses for $\rightarrow$. (Here there is no dependence on any Weak Centering assumption since the $\rightarrow$ construction operates only within worlds.)

Later I will use the following (stated here for the revision-theoretic construction, but with analogs for fixed point): for all worlds $w$, and all ordinals $\alpha$ for (L-i) and all reflection ordinals $\Delta$ for (L-ii):

**(L-i):** If $|A \rightarrow B|_{w,\alpha} = 1$ then $|B \rightarrow C|_{w,\alpha} \leq |A \rightarrow C|_{w,\alpha}$;
**(L-ii):** $|A \rightarrow (B \wedge C)|_{w,\Delta} = \min\{|A \rightarrow B|_{w,\Delta}, |A \rightarrow C|_{w,\Delta}\}$.

The analogs for '$\Rightarrow$' hold as well. Verification of (L-i) is almost trivial. (I'll actually use it only in the case where $\alpha$ is a reflection ordinal, but it holds for all ordinals $\alpha$.) Part of (L-ii) also generalizes to all ordinals:

**(L-iia):** If $|A \rightarrow B|_{w,\alpha} = 1$ then $|A \rightarrow C|_{w,\alpha} \leq |A \rightarrow (B \wedge C)|_{w,\alpha}$

---

[33] The difference between the fixed point constructions for $\rightarrow$ and for $\rhd$ comes in the way that chains of valuations generate valuations: instead of the association given in Appendix B, here when $Z$ is a chain of $\rightarrow$-valuations we use the much simpler:

$$val[Z](w, A \rightarrow B) = \begin{cases} 1 & \text{if } (\exists S \in Z)(\forall u \in S)(|A|_{w,u} = 1 \supset |B|_{w,u} = 1) \\ 0 & \text{if } (\exists S \in Z)(\forall u \in S)(|A|_{w,u} = 1 \wedge |B|_{w,u} = 0) \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

This is basically what's in Section 7 of Field 2014.

(and similarly for $\Rightarrow$), which is likewise easily proved. The remainder of (L-ii) is that when $|A \rightarrow (B \wedge C)|_{w,\Delta} = 0$, one of $|A \rightarrow B|_{w,\Delta}$ and $|A \rightarrow C|_{w,\Delta}$ must be 0. That's so because if $|A \rightarrow B|_{w,\Delta}$ and $|A \rightarrow C|_{w,\Delta}$ are both > 0 then (by the Fundamental Theorem and the evaluation rules) either there's a final $\alpha$ with $|A|_{w,\alpha} < 1$, or both a final $\alpha$ with $|B|_{w,\alpha} > 0$ and a final $\beta$ with $|C|_{w,\beta} > 0$; and then by the Fundamental Theorem again, either $|A|_{w,\Delta} < 1$, or both $|B|_{w,\Delta} > 0$ and $|C|_{w,\Delta} > 0$. So $|A \rightarrow (B \wedge C)|_{w,\Delta+1} > 0$ and (by the Fundamental Theorem once again) $|A \rightarrow (B \wedge C)|_{w,\Delta} > 0$.[34]

**§6. The two types of conditionals together.** So, we know several ways of getting naive truth in a language $L$ with '$\rhd$', and corresponding ways of getting naive truth in a language $L^*$ with '$\rightarrow$'. But what we really want is a language $L^{**}$ with both (and with no restrictions on the embedding of either within the scope of the other).

There are three *prima facie* possible ways to proceed.

The *symmetric option* is to give a single construction (revision or fixed point, as one chooses) that evaluates both kinds of conditionals simultaneously: on the revision approach, this would involve, at each stage $\alpha$, evaluating both $|A \rhd B|_{w,\alpha}$ and $|A \rightarrow B|_{w,\alpha}$ on the basis of the various $|A|_{x,\beta}$ and $|B|_{x,\beta}$ for $\beta < \alpha$ (restricting to the case where $x$ is $w$ in the case of $\rightarrow$).

The $\rhd$*-first option* is to temporarily hold a valuation $v$ for $\rightarrow$ fixed, and use a construction for $\rhd$ on the basis of it. In the case of a revision construction, this would lead, for each choice of $v$, to a reflection ordinal $\Omega_v$ and thus a privileged valuation $j^v$ (= $j^v_{\Omega_v}$) for $\rhd$; in the case of a fixed point construction we similarly get a privileged valuation $j*^v$. Call this the "inner construction". We then would give an "outer" construction (again either revision-theoretic or fixed point; and it needn't be the same choice as for the inner) of a valuation for $\rightarrow$, one that looks only at the privileged valuations of $\rhd$-conditionals constructed in inner constructions from other valuations. For instance, in the case where both the inner and outer constructions are revision-theoretic, we would construct $v_{\alpha+1}$ using valuations of sentences where $\rightarrow$-conditionals are evaluated by $v_\alpha$ and $\rhd$-conditionals by the corresponding $j^{v_\alpha}$ (and use the same rule for limit ordinals as before), eventuating in a reflection ordinal $\Delta$ for the whole construction.

The $\rightarrow$*-first option* is just the reverse. In the case when both inner and outer constructions are revision-theoretic, we temporarily hold fixed a valuation $j$ for $\rhd$, and use a revision construction for $\rightarrow$ on the basis of it; this leads, for each choice of $j$, to a reflection ordinal $\Delta_j$ and thus a privileged valuation $v^j$ (= $v^j_{\Delta_j}$) for $\rightarrow$. That is the "inner construction". We then would give an "outer" construction of a valuation $j$ for $\rhd$, where each $j_{\mu+1}$ is determined from an evaluation of sentences that uses $j_\mu$ and the corresponding $v^{j_\mu}$, eventuating in a reflection ordinal $\Omega$ for the whole construction.

These three choices lead to significantly different results for the joint logic of $\rhd$ and $\rightarrow$. I think the $\rightarrow$-first option is most natural: very roughly, it involves settling the valuation of $\rightarrow$ at each world before doing the $\rhd$-construction which relates different worlds. But the ultimate rationale for the $\rightarrow$-first option is that it leads to by far the most plausible and useful laws of restricted quantification.[35] Some of the laws it leads to will be listed in Section 7. Few of them would hold on either the symmetric or $\rhd$-first options: in the

[34] Had we used the valuation rules for the "Official Conditional" of Field 2008, we would only have gotten (L-iia), not (L-ii).

[35] Field 2014 used fixed point constructions rather than revision constructions for inner and outer, but the decision to take the restricted quantifier conditional as inner was the same there as here. (Recall from note 31 the confusing difference in notation: the restricted quantifier conditional

case of the revision construction, that's because on those options, the validity of a sentence of form $A \rhd B$ (where $A$ and $B$ may contain $\rightarrow$) would require that $B$ has value 1 when $A$ does *at all final ordinals* in the $\rightarrow$-construction, not just at reflection ordinals of the $\rightarrow$-construction. For instance, it's only at reflection ordinals where $A$ and $A \rightarrow \perp$ are prevented from simultaneously having value 1; because of this, the law $[(A \rightarrow B) \wedge A] \rhd B$ couldn't possibly hold on the symmetric or $\rhd$-first options, where it does on the $\rightarrow$-first. (Similar remarks hold for the fixed point constructions.) For more remarks related to this, see note 42 below.

Let's recap (or make explicit) how the overall construction goes on the $\rightarrow$-first option. (I'll stick to the case where both the inner and outer constructions are revision-theoretic.) We start with a 2-valued worlds model $M_0$ for the 'True'-free fragment of $L^{**}$ (whose number-theoretic part is an $\omega$-model in each world, as before). Its ground fragment $L^{**}{}_0$ is to be evaluated either by Burgess 2-valued or variant-Burgess 3-valued semantics. In the former case, '$\rightarrow$' is to be evaluated like '$\supset$' in the ground language. In the latter case, it is to be evaluated in the ground language by the rule that $|A \rightarrow B|$ is 1 whenever $|A| < 1$ and is $|B|$ when $|A| = 1$. (This leads to $\Rightarrow$ being evaluated in the ground language by the 3-valued Lukasiewicz rules: $|A \Rightarrow B|$ is 1 iff $|A| \leq |B|$, 0 iff $|A|$ is 1 and $|B|$ is 0, $\frac{1}{2}$ iff $|A|$ exceeds $|B|$ by $\frac{1}{2}$.) For convenience we expand the language $L^{**}$ by adding names for all objects in the domain $U$ of $M_0$, getting $L^{**+}$.

Now let $T$ be any function that assigns to every object of the ground model a value in $\{0, \frac{1}{2}, 1\}$, subject to the condition that if an object isn't the Gödel number of a sentence of $L^{**}$, $T$ assigns it 0. Let $j$ be any function that assigns to every $L^{**+}$-sentence of form $A \rhd B$ a value in $\{0, \frac{1}{2}, 1\}$, and $\upsilon$ be any function that assigns to every $L^{**+}$-sentence of form $A \rightarrow B$ a value in $\{0, \frac{1}{2}, 1\}$. We now evaluate every $L^{**+}$-sentence relative to $T$, $j$, and $\upsilon$ by essentially the Kleene rules early in Section 4; the only differences are that there is an additional parameter $\upsilon$ in all the valuations, and we have an additional trivial clause for $\upsilon$ analogous to that for $j$:

$$|A \rightarrow B|_{w,j,\upsilon,T} = \upsilon(w, A \rightarrow B).$$

Then, keeping $j$ and $\upsilon$ fixed, we construct the minimal fixed point $T_{min}$ (which now depends on $\upsilon$ as well as on $M_0$ and $j$), and abbreviate $|A|_{w,j,\upsilon,T_{min}}$ as $|A|_{w,j,\upsilon}$.

Next we do the "inner construction": we hold the valuation $j$ for $\rhd$-sentences fixed, and do a revision construction for valuations $\upsilon_\alpha$ of $\rightarrow$-sentences. Adding a subscript $j$ to make explicit the dependence on that $\rhd$-valuation, the stages are given by:

$$|A \rightarrow B|_{w,j,\alpha} = \begin{cases} 1 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta,\alpha))[|A|_{w,j,\gamma} = 1 \supset |B|_{w,j,\gamma} = 1] \\ 0 & \text{if } (\exists\beta < \alpha)(\forall\gamma \in [\beta,\alpha))[|A|_{w,j,\gamma} = 1 \wedge |B|_{w,j,\gamma} = 0] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

For each $j$, we are led to reflection ordinals $\Delta$ (which may depend on $j$ as well as on $M_0$). And the dependence on a $j$ clearly does nothing to block the Fundamental Theorem: we have

PROPOSITION. ***[Fundamental Theorem for** $\rightarrow$ **in** $L^{**}$.] For any $j$, any $j$-reflection ordinal $\Delta$, any $w \in W$, and any $L^{+}$-sentence $A$,*
*(a) $|A|_{w,j,\Delta} = 1$ if and only if $(\forall\gamma \in FIN)(|A|_{w,j,\gamma} = 1)$*
*and (b) $|A_{w,j,\Delta} = 0$ if and only if $(\forall\gamma \in FIN)(|A|_{w,j,\gamma} = 0)$.*

Since there is only one possible value other than 0 and 1, these two clauses imply that each $j$-reflection ordinal $\Delta$ is associated with the same $\rightarrow$-valuation $\upsilon^{\Delta}$; we can call this

there was $\blacktriangleright$, and the $\rightarrow$ there was somewhat in the spirit of the $\rhd$ here.) The inner construction there was called the "fiber construction", and the outer construction the "base space construction".

valuation $v(j)$. Since the particular $\Delta$ doesn't matter as long as it is a $j$-reflection ordinal, we can define $|A|_{w,j}$ to be $|A|_{w,j,\Delta}$ where $\Delta$ is any $j$-reflection ordinal.

In short, for each $j$-valuation for $\rhd$-sentences, we've assigned a privileged valuation $v(j)$ for $\rightarrow$-sentences. (And a minimal fixed point for truth, based on both.) That's the inner construction.

We now use the privileged $v(j)$'s for each $j$ in constructing a specific $j$ for $\rhd$-sentences (the "outer construction"). So unlike in the inner construction, we don't need to add a new parameter $v$ for the valuation of the other conditional $\rightarrow$: the clauses for the $j_\mu$ that evaluate $\rhd$-sentences are EXACTLY as in Section 4.

This may seem to simplify matters, but it actually makes them somewhat more complicated: for the $v$ we use is no longer held constant, it varies with the $j$ in the revision process. Because of this, we need to revisit the Fundamental Theorem for $\rhd$: in particular, the induction on complexity in Stage (2) of the proof in Appendix A. For now we must consider in the induction step not only sentences of form $\neg B$, $B \wedge C$, $\forall x B$ and $\Box A$, but also sentences of form $A \rightarrow B$. And it's unobvious how to carry out the induction step in this case.

Indeed, it's more than unobvious: it can't be done, **the Fundamental Theorem for $\rhd$ no longer holds without restriction once $\rightarrow$ is added to the language**. Example: As a preliminary, let $K_\rhd$ be constructed (by the usual Gödel-Tarski procedure) to be equivalent to $True(\langle K_\rhd \rangle) \rhd \neg True(\langle K_\rhd \rangle)$, and hence given naivety to $K_\rhd \rhd \neg K_\rhd$. On the semantics as given, at each world $w$ for which $W_w \neq \emptyset$ (which includes all those $w$ at which there is at least Weak Centering) and for each stage $\kappa$ for the outer construction,

**(\$):** $|K_\rhd|_{w,\kappa}$ is 1 if $\kappa$ is odd, 0 if $\kappa$ is an even successor, and $\frac{1}{2}$ if $\kappa$ is a limit.

(That's so both for the semantics based on the original Burgess and the one based on the variant.) Now let $K^*$ be $K_\rhd \rightarrow \neg K_\rhd$. Since $K_\rhd$ is equivalent to a $\rhd$-conditional, its value is held fixed during any $\rightarrow$-construction, so at each $w$ and each stage $\kappa$ for the outer construction and each stage $\alpha$ >0 for the inner, $|K^*|_{w,\kappa,\alpha}$ is 1 if $|K_\rhd|_{w,\kappa} < 1$ and is 0 otherwise. So using (\$), when $W_w \neq \emptyset$, $|K^*|_{w,\kappa}$ (i.e. $|K^*|_{w,\kappa,\Delta}$) is 1 when $\kappa$ is even (including when it is a limit), and 0 otherwise. So for any world, at $\kappa = \Omega$, $K^*$ has value 1, but not at all final $\kappa$ in the $\rhd$-construction.[36]

The failure of the Fundamental Theorem for $\rhd$ is not devastating, for we still get the special case of it for $\rhd$-conditionals, which is what is needed for many laws, such as Modus Ponens (assuming Weak Centering for $\rhd$). Indeed, we get more generally that the Fundamental Theorem for $\rhd$ holds for every sentence $A$ in which all occurrences of 'True' and '$\rightarrow$' are inside the scope of an '$\rhd$'.

The special case of the Fundamental Theorem for $\rhd$ is enough to establish that all reflection ordinals in the $j_v$ construction give rise to the same values for every sentence: for it immediately gives this for every $\rhd$-conditional, and the generalization to all sentences is immediate by induction.

## §7. Application to restricted quantification.

Here are some highly desirable laws of restricted quantification: it is hard to imagine making serious use of restricted quantification without them, or at least, something very close to them. Indeed, we should expect them

---

[36] It won't help to alter the starting point of the $\rightarrow$-construction, e.g. by making conditionals start with value $\frac{1}{2}$ at some worlds but 1 at some and 0 at others. There are several reasons, but the main one is that the evaluation of $K_\rhd$ would even out by stage $\omega$, so that (\$) would still hold for infinite $\kappa$.

to be strong laws in the sense explained in Section 4.1, which guarantees that prefixing any string of $\Box$s and $\Diamond$s to one of them is also to be a law, and that they remain valid however their antecedent is strengthened.[37] (The four with an asterisk are obtained using $\rhd$-contraposition from their unasterisked counterparts;[38] but since $\rhd$-contraposition isn't generally valid for variably strict conditionals they need to be stated separately. The ones marked 'b' result from the corresponding ones marked 'a' by a kind of quasi-contraposition which is also not generally valid for variably strict conditionals.) I've written these laws with $\Rightarrow$, reflecting my view that the conditional for restricted quantification is contraposable, but until we get to CQ, every law on the list would remain valid were $\Rightarrow$ to be replaced with $\rightarrow$.

**1:** $[\forall x(Ax \Rightarrow Bx) \wedge Ay] \rhd By$ "If all $A$ are $B$, and $y$ is $A$, then $y$ is $B$"

**2:** $\forall x Bx \rhd \forall x(Ax \Rightarrow Bx)$ "If everything is $B$, then all $A$ are $B$"

**2*:** $\neg\forall x(Ax \Rightarrow Bx) \rhd \neg\forall x Bx$ "If not all $A$ are $B$, then not everything is $B$"

**3a:** $\forall x(Ax \Rightarrow Bx) \wedge \forall x(Bx \Rightarrow Cx) \rhd \forall x(Ax \Rightarrow Cx)$
"If all $A$ are $B$ and all $B$ are $C$ then all $A$ are $C$"

**3b:** $\forall x(Ax \Rightarrow Bx) \wedge \neg\forall x(Ax \Rightarrow Cx) \rhd \neg\forall x(Bx \Rightarrow Cx)$
"If all $A$ are $B$ and not all $A$ are $C$ then not all $B$ are $C$"

**4a:** $\forall x(Ax \Rightarrow Bx) \wedge \forall x(Ax \Rightarrow Cx) \rhd \forall x(Ax \Rightarrow Bx \wedge Cx)$
"If all $A$ are $B$ and all $A$ are $C$ then all $A$ are both $B$ and $C$"

**4b:** $\forall x(Ax \Rightarrow Bx) \wedge \neg\forall x(Ax \Rightarrow Bx \wedge Cx) \rhd \neg\forall x(Ax \Rightarrow Cx)$
"If all $A$ are $B$ and not all $A$ are both $B$ and $C$ then not all $A$ are $C$"

**4a*:** $\neg\forall x(Ax \Rightarrow Bx \wedge Cx) \rhd \neg\forall x(Ax \Rightarrow Bx) \vee \neg\forall x(Ax \Rightarrow Cx)$
"If not all $A$ are both $B$ and $C$ then either not all $A$ are $B$ or not all $A$ are $C$"

**5:** $\neg\forall x(Ax \Rightarrow Bx) \rhd \exists x(Ax \wedge \neg Bx)$
"If not all $A$ are $B$, then something is both $A$ and not $B$"

**5*:** $\forall x(\neg Ax \vee Bx) \rhd \forall x(Ax \Rightarrow Bx)$
"If everything is either not-$A$ or $B$, then all $A$ are $B$" / "If nothing is both $A$ and not-$B$, then all $A$ are $B$"

**6:** $\exists x(Ax \wedge \neg Bx) \rhd \neg\forall x(Ax \Rightarrow Bx)$
"If something is both $A$ and not $B$, then not all $A$ are $B$"

**CQ:** $\forall x(Ax \Rightarrow Bx) \rhd \forall x(\neg Bx \Rightarrow \neg Ax)$ "If all $A$ are $B$ then all not-$B$ are not-$A$".

**CQ*:** $\neg\forall x(Ax \Rightarrow Bx) \rhd \neg\forall x(\neg Bx \Rightarrow \neg Ax)$
"If not all $A$ are $B$ then not all not-$B$ are not-$A$".

(There is a bit of redundancy in the list: 2* follows by obvious laws from 5, and 2 from 5*.)

CQ and CQ* strike me as less *obviously* desirable than the earlier members of the list. However, CQ together with 1 and 2 respectively (and double negation laws in the case of 2) yield:

**1c:** $[\forall x(Cx \Rightarrow Dx) \wedge \neg Dy] \rhd \neg Cy$ "If all $C$ are $D$, and $y$ is not $D$, then $y$ is not $C$"

**2c:** $\forall x \neg Cx \rhd \forall x(Cx \Rightarrow Dx)$ "If nothing is $C$, then all $C$ are $D$"

And these do seem to me obviously desirable; indeed, no less so than the laws 1 and 2 from which they were obtained. It's unobvious how to get a plausible theory that delivers 1c and 2c without delivering CQ (and probably CQ*), which I take to provide support for the latter.

[37] Note that though the proof of the latter in note 26 relied on the Fundamental Theorem, it used it only for $\rhd$-sentences, so it still holds when $\rightarrow$ is in the language.

[38] With double negation laws (and re-lettering) in the case of CQ*.

Still, someone willing to give up 1c and 2c could use the results of this paper to validate the laws of restricted quantification preceding CQ with a restricted quantifier based on $\rightarrow$ instead of $\Rightarrow$.

Despite the desirability of these laws, it is not entirely easy to give an account of conditionals in naive truth theory that validate them all (even without the modal prefixes). Indeed, prior to Field 2014, no published theory came close. But there are two precursors worth mentioning, Beall *et al* 2006 and Beall 2009. Both are in a paraconsistent framework, which means (given reasonable assumptions that they accept) that they can't accept a restricted-quantifier analog of law 2c, or even of its rule form. For if $Cx$ means $x = x \wedge A$ and $Dx$ means $x = x \wedge B$ then even the rule version of 2c requires that $\neg A$ imply $A \mapsto B$ (where $\mapsto$ is the paraconsistent restricted quantifier conditional); and then Modus Ponens yields Explosion. To deal with this, both precursors propose that the conditional that restricts quantification be non-contraposable,[39] i.e. they disallow even the rule form of CQ for $\mapsto$ (and CQ*, given previous note). Myself, I'm not happy with the loss of 2c; but neither account does well with other laws either.

Beall *et al* 2006 made an important contribution in focusing on the need of a logic of restricted quantification and introducing the idea of using two separate conditionals for it. The paper didn't show, or even claim, that a naive truth theory could be added without triviality to the main logics it considers (those in their Section 6); but their discussion is explicitly motivated by the hope/belief that this is so. (One of the authors explicitly stated several years later that the question of non-triviality was open: see Beall 2009, p. 121.) Putting any worries about lack of non-triviality proof aside, the main issue is over the laws. The good news is that their framework validates their analogues of laws 2 and 4a (taking the analogues to have their noncontraposable $\mapsto$ in place of my contraposable $\Rightarrow$, as well as their relevance conditional in place of my $\rhd$); hence also 2* and 4a*, assuming the interpretation in note 39. The bad news is that it doesn't validate any of the others (though it does validate rule forms of some of them). Also, the validation of 2 and 2* depends very directly on their assumption of the validity of

**(?):** $A \rhd B \models A \mapsto B$.

And (?) immediately rules out the analog of my law 1 (when naive truth, Modus Ponens for $\mapsto$, and reasonable quantifier laws are present). The reason is that given reasonable quantifier laws, law 1 requires $[(A \mapsto B) \wedge A] \rhd B$; and then (?) delivers

**Pseodo Modus Ponens:** $[(A \mapsto B) \wedge A] \mapsto B$.

And it's well-known that this is inconsistent with genuine Modus Ponens for $\mapsto$ (i.e. $(A \mapsto B) \wedge A \models B$) in a naive theory (assuming the standard structural rules for validity mentioned in note 3).[40] The centrality of (?) to the derivation of law 2 suggests that no simple modification of the account is likely to yield laws 1 and 2 together.

[39] Interestingly, they take their main conditional (a relevance conditional, their analog of my $\rhd$) to obey a rule form of contraposition. (Beall 2009 very clearly does; Beall *et al* 2006 is slightly equivocal: see p. 595 middle.) I take this to mean that their main conditional isn't a good candidate for an account of the ordinary indicative conditional: see the Trump example in Section 2.

[40] In the logic I've been advocating (with Weak Centering assumed so as to get Modus Ponens), we do have

$$C \vee \neg C, C \rhd B \models C \Rightarrow B;$$

but the need for the excluded middle premise is sufficient to prevent the paradox.

ZU064-05-FPR S1755020315000301 19 September 2015 12:42



The second precursor is Beall 2009 (pp. 119-226). It also used two separate conditionals for the logic of restricted quantification. It suggests three different options for the logic, and unlike Beall *et al* 2006, shows each to be compatible with naive truth. All of them validate (?), so again it is immediate that law 1 can't be satisfied. The situation for laws is slightly worse than Beall *et al* 2006. Beall's first two options validate only 4a and 4a* from the list (though the weaker rule forms of some of the others are validated). His third option validates only 2 and 2*; indeed, its method of achieving 2 and 2* causes it to violate even the rule form of 4a.

Without going into detail, the main problem in both Beall *et al* 2006 and Beall 2009 arises because (a) a certain kind of "abnormal" worlds are essential to these accounts (unlike the present account, where they are optional); (b) at these worlds, both conditionals are very badly behaved; and (c) the validity of $X \triangleright Y$ (using my notation for their relevance conditional) requires that it be true at all normal worlds, which in turn requires that at all worlds including abnormal ones, $Y$ is true when $X$ is. Collectively these make it very hard for reasonable $\triangleright$-statements with $\mapsto$-conditionals in their antecedents or consequents to come out valid. (An additional problem arises because of the way that these accounts handle negation, via a shift in worlds: this immediately rules out laws like 3b and 4b.)

Field 2014 used a very different framework, and did manage to validate the entire list; but the semantics it employed for $\triangleright$ seemed *ad hoc*. (That paper did note some commonalities between its $\triangleright$ and the ordinary indicative conditional, but also pointed out that the conditional reduced to the material conditional rather than the indicative conditional in 'True'-free contexts.)[41]

But I now note that the entire list is also validated on the semantics of the present paper, with its independently motivated $\triangleright$. (We also get Modus Ponens for $\triangleright$, if we insist on Weak Centering at normal worlds in the base model, as I think we clearly should.)

The real work in establishing the laws on the list has nothing to do with the quantifiers, it's all in the relation among conditionals. The laws we need are the results of prefixing the following with strings of $\Box$s and $\Diamond$s:

**I:** $[(A \Rightarrow B) \wedge A] \triangleright B$ (for 1)
**IIIa:** $(A \Rightarrow B) \wedge (B \Rightarrow C) \triangleright (A \Rightarrow C)$ (for 3a)
**IIIb:** $(A \Rightarrow B) \wedge \neg(A \Rightarrow C) \triangleright \neg(B \Rightarrow C)$ (for 3b)
**IVa:** $(A \Rightarrow B) \wedge (A \Rightarrow C) \triangleright (A \Rightarrow B \wedge C)$ (for 4a)
**IVb:** $(A \Rightarrow B) \wedge \neg(A \Rightarrow B \wedge C) \triangleright \neg(A \Rightarrow C)$ (for 4b)
**IVa*:** $\neg(A \Rightarrow B \wedge C) \triangleright \neg(A \Rightarrow B) \vee \neg(A \Rightarrow C)$ (for 4a*)
**V:** $\neg(A \Rightarrow B) \triangleleft\triangleright (A \wedge \neg B)$ (for 5, 2* and 6)
**V*:** $(\neg A \vee B) \triangleright (A \Rightarrow B)$ (for 5* and 2)
**C:** $(A \Rightarrow B) \triangleleft\triangleright (\neg B \Rightarrow \neg A)$ (for CQ)
**C*:** $\neg(A \Rightarrow B) \triangleleft\triangleright \neg(\neg B \Rightarrow \neg A)$ (for CQ*)

41 Despite its reducing to the material conditional, we can in retrospect see the conditional of Field 2014 as pretty much a degenerate case of the indicative conditional of the present paper. For the construction there started from a classical first order model, which can be seen as a degenerate Burgess model with only one world, weakly centered (which in the one-world case means simply "accessible from itself"). In that degenerate case, '$\triangleright$' obviously coincides with the material conditional in the ground model. (The conditional there still differed in a small respect from the degenerate case of the current construction: it utilized what I there called "dynamic Kripke constructions". I have dropped them here since they don't yield the results that we want once we clearly focus on extending the *ordinary indicative* conditional to a language with 'True'.)

ZU064-05-FPR S1755020315000301 19 September 2015 12:42



C and C* are of course entirely trivial given the definition of $\Rightarrow$ in terms of $\rightarrow$. For most of the others, the proof is almost immediate from what has already been said, especially at the end of Section 5. (The analogs of these latter laws for $\rightarrow$ hold equally.) For note that to establish that a claim of form $P(X \rhd Y)$ is valid, where $P$ is any string of $\Box$s and $\Diamond$s, it suffices to show (in the revision-theoretic version; but the fixed point is analogous) that for all worlds $w$ and all final $\kappa$ of the $\rhd$-construction, if $|X|_{w,\kappa} = 1$ then $|Y|_{w,\kappa} = 1$. In other words, that for all $w$ and $\kappa$, and all $\kappa$-reflection ordinals $\Delta_\kappa$ of the $\rightarrow$-construction, if $|X|_{w,\kappa,\Delta_\kappa} = 1$ then $|Y|_{w,\kappa,\Delta_\kappa} = 1$. Given this, the proof of I is immediate from "Modus Ponens for $\rightarrow$ and $\Rightarrow$", and V from "Weak equivalence of $\neg(A \rightarrow B)$ and $\neg(A \Rightarrow B)$ to $A \wedge \neg B$". And IIIa and IIIb follow from the special case of (L-i) (end of Section 5) where $\alpha$ is $\Delta$, and IVa, IVb and IVa* from (L-ii). As for V*, if $|\neg A \vee B|_{w,\kappa,\Delta_\kappa} = 1$ then either $|A|_{w,\kappa,\Delta_\kappa} = 0$ or $|B|_{w,\kappa,\Delta_\kappa} = 1$, and so by the Fundamental Theorem for $\rightarrow$, either for all $\kappa$-final $\alpha$, $|A|_{w,\kappa,\alpha} = 0$ or else for all $\kappa$-final $\alpha$, $|B|_{w,\kappa,\alpha} = 1$; in either case, for all $\kappa$-final $\alpha$, $|A|_{w,\kappa,\alpha} \leq |B|_{w,\kappa,\alpha}$. From this it clearly follows that for all final $\alpha$ $|A \Rightarrow B|_{w,\kappa,\alpha} = 1$ and hence in particular that $|A \Rightarrow B|_{w,\kappa,\Delta_\kappa} = 1$.

This only scratches the surface of the logic of the system,[42] but it is not my purpose here to explore it at all systematically: my purpose is simply to show that it does easily lead to obvious laws of restricted quantification, which other approaches to conditionals in naive truth theory (other than the *ad hoc* one of Field 2014) haven't come close to meeting. And I think that by basing the laws on an independently motivated account of indicative conditionals, the resulting theory is quite natural.

In particular, it's worth emphasizing that the use of two distinct conditional operators (which is essential for the compatibility of the logic with naive truth, since if $\rightarrow$ and $\rhd$ were identified then we'd have the disastrous (?)) is independently motivated: as I argue at the beginning of Section 5, we can see independently of the laws recently listed that the indicative conditional and the conditional for restricted quantification must be different.

**Thanks:** Harvey Lederman, Graham Priest and two anonymous referees made comments that have led to significant improvements.

**Appendix A: Proof of Fundamental Theorem for $L$ (revision-theoretic version).** Since $\Omega \in FIN$, the right to left of (a) and (b) in the Theorem are trivial. Contraposing the left to right and making the Kripke-stages $\sigma$ explicit, what we need to establish is that for any reflection ordinal $\Omega$ and any $L^+$-sentence $A$:

(a*) $(\forall w \in W)$[if $(\exists \nu \in FIN)(|A|_{w,\nu} < 1)$ then $\forall\sigma(|A|_{w,\Omega,\sigma} < 1)$], and

(b*) $(\forall w \in W)$[if $(\exists \nu \in FIN)(|A|_{w,\nu} > 0)$ then $\forall\sigma(|A|_{w,\Omega,\sigma} > 0)$].

---

42 The reader will note that the schemas I've listed and proved are ones where there are no occurrences of $\rhd$ inside the scope of an $\rightarrow$ (or an $\Rightarrow$). This is no accident: the $\rightarrow$-first construction makes it much easier for a schema in which $\rightarrow$ is in the scope of $\rhd$ to be valid than for one where $\rhd$ is in the scope of $\rightarrow$ to be valid. I think that schemas of the latter sort tend to be far less important than the former (recall the frequently-voiced claim that embeddings of indicative conditionals in the scope of other operators are hard to interpret); that is the main reason I went for an $\rightarrow$-first option. (I have however made no prohibitions on the well-formedness of embeddings of $\rhd$ inside the scope of $\rightarrow$; and with any valid schema such as those listed, there are instances of the schema with arbitrarily complex chains of embeddings of $\rhd$ and $\rightarrow$.)

Despite what I've just said, there are important laws that depend on the embedding of '$\rhd$' in the scope of '$\rightarrow$', but these are mostly meta-rules, whose legitimacy is not blocked by the $\rightarrow$-first option. A typical example is the meta-rule stated in note 28, whose proof is routine.

ZU064-05-FPR S1755020315000301 19 September 2015 12:42



We establish (a*) and (b*) in three steps:

(1) In the special case when $A$ is a conditional $B \rhd C$, the value of $\sigma$ makes no difference, and by the fact that $\Omega$ is a limit ordinal and the evaluation rules for conditionals are continuous with respect to 1 and 0 at limits, the claims are just:

(a*-s) $(\forall w \in W)$[if $(\exists \nu \in FIN)(j_{w,\nu}(B \rhd C) < 1)$ then $(\forall \mu < \Omega)(\exists \nu \in [\mu, \Omega))$ $[j_{w,\nu}(B \rhd C) < 1)]$;

(b*-s) $(\forall w \in W)$[if $(\exists \nu \in FIN)(j_{w,\nu}(B \rhd C) > 0)$ then $(\forall \mu < \Omega)(\exists \nu \in [\mu, \Omega))$ $[j_{w,\nu}(B \rhd C) > 0)]$.

But by (ii) of [Gupta-Belnap] these are trivial.

(2) Given (1), we can show (for any $\sigma$) that if (a*) and (b*) hold for the special case where $A$ is of form '$True(c)$' when $c$ denotes the Gödel number of a sentence, then they hold for all $L^+$-sentences $A$. This is a routine induction on complexity, counting $\rhd$-sentences as of complexity 0 for the purposes of the induction: the claim is trivial for all other atomic sentences since they keep the same value at every revision-stage $\nu$, and the induction step for sentences $\neg B$, $B \wedge C$, $\forall x B$ and $\Box B$ is easy. For instance for $\Box$: (a) Suppose that for some world $w$, $(\exists \nu \in FIN)(|\Box B|_{w,\nu} < 1)$. Then $(\exists \nu \in FIN)$ $(\exists y \in W_w)(|B|_{y,\nu} < 1)$; reversing the quantifier order and applying the induction hypothesis, we get that for some $y \in W_w$, $|B|_{y,\Omega,\sigma} < 1$ (for any $\sigma$), and so $|\Box B|_{w,\Omega,\sigma} < 1$. (b) Suppose that for some world $w$, $(\exists \nu \in FIN)(|\Box B|_{w,\nu} > 0)$. Then $(\exists \nu \in FIN)$ $(\forall y \in W_w)(|B|_{y,\nu} > 0)$; so certainly for all $y$ in $W_w$, $(\exists \nu \in FIN)(|B|_{y,\nu} > 0)$, and by the induction hypothesis for all $y$ in $W_w$, $|B|_{y,\Omega,\sigma} > 0$ (for any $\sigma$); so $|\Box B|_{w,\Omega,\sigma} > 0$ for any $\sigma$.

(3) It remains only to show that for all Kripke-stages $\sigma$ and all $c$ that denote Gödel numbers of sentences, (a*) and (b*) hold for sentences of form '$True(c)$'. But this is trivial when $\sigma = 0$, since $|True(c)|_{w,\Omega,0}$ is always $\frac{1}{2}$ . We now show that if it holds for $\sigma = \tau$ then it holds for $\sigma = \tau + 1$. Suppose $c$ denotes $B$. Then by the assumption about $\tau$ and the result (2), we get

$(\forall w \in W)$[if $(\exists \nu \in FIN)(|B|_{w,\nu} < 1)$ then $|B|_{w,\Omega,\tau} < 1]$

and the analog with '$> 0$' instead of '$< 1$'; which by the transparency of the $j_\nu$-valuations and the Kripke construction gives

$(\forall w \in W)$[if $(\exists \nu \in FIN)(|True(c)|_{w,\nu} < 1)$ then $|B|_{w,\Omega,\tau} < 1]$

and its analog. But by the valuation rules, $|B|_{w,\Omega,\tau}$ is the same as $|True(c)|_{w,\Omega,\tau+1}$, so the result is established. The case where $\sigma$ is a limit ordinal is trivial: no sentence of form '$True(c)$' first passes from $\frac{1}{2}$ to another value at a limit stage of the Kripke construction.

**Appendix B: The fixed point construction for $L$.** Again, a valuation function is a function that assigns to each world and $L^+$-conditional a value in $\{0, \frac{1}{2}, 1\}$.

Let a *chain* be a set $P$ of nonempty sets of transparent valuation functions, meeting the condition that if $S_1, S_2 \in P$ then either $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$.

Given a chain $P$, define a valuation function $val[P]$ ("the valuation function generated by $P$") as:

$$val[P](w, A \rhd B) \text{ is } \begin{cases} 1 & \text{if } (\exists S \in P)(\forall j \in S)(\forall x \in W_w)[|A|_{x,j} = 1 \supset (\exists y \leq_w x) \\ & \quad [|A|_{y,j} = 1 \wedge (\forall z \leq_w y)(|A|_{z,j} = 1 \supset |B|_{z,j} = 1)]] \\ 0 & \text{if } (\exists S \in P)(\forall j \in S)[(\forall x \in W_w)[|A|_{x,j} = 1 \supset (\exists y \leq_w x) \\ & \quad [|A|_{y,j} = 1 \wedge (\forall z \leq_w y)(|A|_{z,j} = 1 \supset |B|_{z,j} = 0)] \wedge \\ & \quad (\exists x \in W_w)(|A|_{x,j} = 1)]] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

(This is for the semantics based on modified Burgess; for that based on original Burgess, the modification of the 0 clause is obvious.) Clearly each $val[P]$ is transparent, given that members of $\cup P$ are.

Let $P_1 \leq P_2$ mean that every member of $P_1$ has a subset that's a member of $P_2$. (Having small members makes a chain bigger.) Chains that are smaller in this ordering generate weaker valuation functions: if $P_1 \leq P_2$ then for all $w$, $val_w[P] \leq_K val_w[P]$. (That's simply because the 1 clause and 0 clause both have form "$(\exists S \in P)(\forall j \in S)$...".)

Define a sequence $J_\mu$ of sets of transparent valuation functions:

$$J_\mu = \{val[P] : P \text{ is a chain and } (\forall \beta < \mu)(\exists S \in P)(S \subseteq J_\beta)\}.$$

For $\mu > 0$ an equivalent and perhaps more intuitive definition of $J_\mu$ is: $\{val[P] : P$ is a non-empty chain and $(\forall \beta < \mu)(\forall S \in P)(S \subseteq J_\beta)\}$. This is more restrictive about the chains, but it's easy to see that any valuation generated by one of the chains in the original is generated also by one of the more restrictive ones.

If $\mu < \nu$, $J_\nu \subseteq J_\mu$, so obviously we eventually reach a fixed point $\mathbf{J}$. That would be uninteresting if $\mathbf{J}$ were empty, but it can be shown (following the model of Field 2014, section 7) that $\mathbf{J} \neq \emptyset$. So letting $\mathbf{P}$ be the set of $\mathbf{J}$-chains (chains whose members are all subsets of $\mathbf{J}$) we'll have

**(FP):** $\mathbf{J} = \{val[P] : P \in \mathbf{P}\}$.

This sets up a one-many correspondence between the $j$ in $\mathbf{J}$ and the $P$ in $\mathbf{P}$. (The members of $\mathbf{J}$ are the analogs in this construction of the valuation functions associated with ordinals in $FIN$ in the revision construction.)

The $\leq$-minimal chain is $\{\mathbf{J}\}$; let $j^*$ be the valuation it generates, i.e. $val[\{\mathbf{J}\}]$. This is the analog, in the fixed point construction, of the valuation function at reflection ordinals. We have

$$|A \rhd B|_{w,j^*} = \begin{cases} 1 & \text{if } (\forall j \in \mathbf{J})(\forall x \in W_w)[|A|_{x,j} = 1 \supset (\exists y \leq_w x) \\ & [|A|_{y,j} = 1 \wedge (\forall z \leq_w y)(|A|_{z,j} = 1 \supset |B|_{z,j} = 1)]] \\ 0 & \text{if } (\forall j \in \mathbf{J})[(\forall x \in W_w)[|A|_{x,j} = 1 \supset (\exists y \leq_w x) \\ & [|A|_{y,j} = 1 \wedge (\forall z \leq_w y)(|A|_{z,j} = 1 \supset |B|_{z,j} = 0)]] \wedge \\ & (\exists x \in W_w)(|A|_{x,j} = 1)] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

In this case the Fundamental Theorem, as stated in the text, concerns the special nature of the valuation function at $j^*$. Its proof and the proof of the fixed point result are a simple adaptation of that in Section 7 of Field 2014.

## BIBLIOGRAPHY

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

Beall, J. C., Brady, R., Hazen, A., Priest, G., & Restall, G. (2006). Relevant Restricted Quantification. *Journal of Philosophical Logic*, **35**, 587–598.

Burgess, J. (1981). Quick Completeness Proofs for some Logics of Conditionals. *Notre Dame Journal of Formal Logic*, **22**, 76–84.

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

Field, H. (2014). Naive Truth and Restricted quantification: Saving Truth a Whole Lot Better. *Review of Symbolic Logic*, **7**, 147–91.

Gupta, A. and Belnap, N. (1993). *The Revision Theory of Truth*. Cambridge, MA: MIT Press.

Kripke, S. (1975). Outline of a Theory of Truth. *Journal of Philosophy*, **72**, 690–716.

Lewis, D. (1974). *Counterfactuals*. Cambridge, MA: Harvard University Press.

McGee, V. (1985). A Counterexample to Modus Ponens. *Journal of Philosophy*, **82**, 462–71.

Pollock, J. (1976). The Possible Worlds Analysis of Counterfactuals. *Philosophical Studies*, **29**, 469–76.

Restall, G. (2007). Curry's Revenge: The Costs of Non-Classical Solutions to the Paradoxes of Self-Reference. In Beall J. C., editor. *Revenge of the Liar*, Oxford: Oxford University Press.

Stalnaker, R. (1968). A Theory of Conditionals. In Rescher, N., editor. *Studies in Logical Theory*, Oxford: Blackwell Publishers.

Yablo, S. (2003). New Grounds for Naive Truth Theory. In Beall, J. C., editor, *Liars and Heaps*, Oxford: Oxford University Press.

DEPARTMENT OF PHILOSOPHY
NEW YORK UNIVERSITY
5 WASHINGTON PLACE
NEW YORK, NY 10003
and
DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF BIRMINGHAM
EDGBASTON, BIRMINGHAM
B15 2TT UK
*Email*: hartryfield@nyu.edu

# PROPERTIES, PROPOSITIONS AND CONDITIONALS

HARTRY FIELD

Abstract. Section 1 discusses properties and propositions, and some of the motivation for an account in which property instantiation and propositional truth behave "naively". Section 2 generalizes a standard Kripke construction for naive properties and propositions, in a language with modal operators but no conditionals. Whereas Kripke uses a 3-valued value space, the generalized account allows for a broad array of value spaces, including the unit interval [0,1]. This is put to use in Section 3, where I add to the language a conditional suitable for restricting quantification. The shift from a value space based on the "mini-space" $\{0, \frac{1}{2}, 1\}$ to one based on the "mini-space" [0,1] leads to more satisfactory results than I was able to achieve in previous work: a vast variety of paradoxical sentences can now be treated very simply. In Section 4 I make a further addition to the language, a conditional modeled on the ordinary English conditional, paying particular attention to how it interacts with the restricted quantifier conditional. This is all done in the [0,1] framework, and two alternatives are considered for how the ordinary conditional is to be handled; one of them results from adding a tweak to a construction by Ross Brady. Section 5 discusses a further alternative, a standard relevance conditional (for the ordinary conditional, perhaps for use with a different quantifier-restricting conditional), but argues that it is not promising. Section 6 discusses the identity conditions of properties and propositions (again in the setting of a value space based on [0,1]); the issue of achieving naivety for coarse-grained properties is seen to be more complicated than some brief remarks in Field 2010 suggested, but a way to get a fair degree of coarse-grainedness is shown.

## 1. Naivety in a Theory of Properties and Propositions

### 1.1. Properties and Propositions.

**1.1. Properties and Propositions.** I take it that the main point of talking about propositions and properties[1] is to provide a natural framework for talking about language and the mind. In the language case, this goes as follows:

**(EXP):** A sentence (as used on a given occasion) is true (at possible or impossible world $w$) if and only if it expresses (on that occasion) a proposition that is true (at $w$).
A formula (as used on a given occasion) is true of something $o$ (at world $w$) if and only if it expresses (on that occasion) a property that is instantiated (at $w$) by $o$.

*Linguistic* truth and satisfaction (the converse of truth-of) are reduced to *propositional* truth and property-instantiation, via the expression relation. In the mental case, e.g. of beliefs, it's similar.

[1]In the "abundant" sense in which there are highly "unnatural" properties like being either hairy or purple or a small ocelet or an African country. A "sparse" conception of properties, e.g. one confined to basic physical properties, might well have another point, but such a conception will not be under discussion.

But what are propositions and properties? And what is it for a proposition to be true (at a possible or impossible world $w$)? And for a property to be instantiated by an object (at such a world)?

A common procedure is to stipulate at the start what propositions and properties are, characterizing them independently of language. For instance, we might stipulate that propositions are just sets of worlds (perhaps including impossible ones), and that properties are just sets of pairs of worlds and things. On this approach, the theory of propositional truth and property-instantiation is trivial: a proposition is true at a world if and only if the world is a member of it, and something $o$ instantiates a property at a world $w$ if and only if the pair $\langle w, o\rangle$ is a member of that property.

Unfortunately, this leaves all the interesting questions unresolved. Propositions and properties are only of use if we can relate them to our sentences and predicates—that is, if we can sensibly speak of the proposition or property that a given one of our sentences or predicates *expresses (on a given occasion of use).* What we're ultimately interested in is what it is for *whatever proposition is expressed (on a given occasion) by 'Snow is white'* to be true at a world—and analogously for sentences other than 'Snow is white', including potentially paradoxical ones like 'Nothing that Joe is saying is true'. Similarly, what we're ultimately interested in is what it is for *whatever property is expressed (on a given occasion) by 'is a red ball'* to be instantiated by a given thing at a world—and analogously for 'is either a red ball or a property that doesn't instantiate itself'.

Moreover, when it comes to *prima facie* paradoxical properties and propositions, the possible worlds definition of propositions and properties begs questions against many approaches. Perhaps there is *some generalization* of the idea of sets of worlds, or sets of world-object pairs, that could be used instead; but in advance of developing the theory it's hard to know whether this is so, and if it is so then what sort of entities will do the trick.[2]

A better approach to the issues of propositions and their truth, and properties and their instantiation, is to start from the idea that our grasp of the notions of proposition and property is based on locutions like "the proposition *that* ___" and "the property *of being* ___", where in the blanks go sentences and formulas of our language that we already understand. I'll introduce abstraction terms for this: $\lambda B$ for "the proposition that $B$" where $B$ is a sentence of our language, $\lambda x Px$ for "the property that $P$" where $Px$ is a formula of our language involving only $x$ free. Then the questions I will start from (before I generalize them below, to allow for parameters) are:

> Under what circumstances is there a proposition $\lambda B$, and when there is, under what circumstances is it true at a given world?

and

[2]There are also technical issues about the worlds account: taken literally, it assumes that the worlds form a set (which in a standard set theory entails that there is a limitation on the size of worlds), and that within each world the things that instantiate a given property form a set (which in a standard set theory rules out for instance that there be a property that at some world is instantiated by all sets). I don't take these to be integral to the spirit of the proposal; the generalization I'm alluding to will avoid these features, and can do so by offering axioms on propositions and propertes rather than a reductive definition.

> Under what circumstances is there a property $\lambda x P$, and when there is, under what circumstances is it instantiated at a given world by a given object?

('Object' here and in what follows is to be construed broadly: *anything*, including properties and propositions, will count as objects.) It will simplify what follows to imagine that our own language is regimented to exclude ambiguous terms or constructions, indexical elements, and the like; otherwise the ambiguities or indexicalities would be inherited by the $\lambda$-terms. Let $L$ be such a regimented version of our language.

I will leave metaphysical questions about the nature of properties and propositions unsettled. But I'm inclined to think that there is nothing to say about properties and propositions beyond the answers to these questions, plus the question of the identity-conditions of properties and propositions to be considered later.

1.2. **Naivety.** The focus of this paper will be on the *naive* theory of properties and propositions, which answers the above questions as follows:

> For each formula $P(x)$ of $L$ with exactly one free variable $x$,
> ($i$) there is a corresponding property $\lambda x P(x)$, and
> ($ii$) $\lambda x P(x)$ is instantiated (at a possible or impossible world $w$)
> by all and only the objects $o$ such that (at $w$) $P(o)$.

Similarly:

> For each sentence $B$ of $L$,
> ($i_0$) there is a corresponding proposition $\lambda B$, and
> ($ii_0$) $\lambda B$ is true (at a world $w$) if and only if (at $w$) $B$.

Such a theory is apparently threatened by paradoxes, e.g. the analog for properties of Russell's paradox for sets; but there is now a large body of literature on ways to keep the naive theory, by weakening the classical logical assumptions used in the derivations of absurdities from it. This paper will suggest some improvements in current approaches to this, with some attention to the assessment of alternatives.

While I won't seriously discuss alternatives to naivety here, I will make one remark regarding the most obvious alternative: a *non-existence theory*, which for paradoxical predicates denies (i), but which accepts that (ii) holds whenever $\lambda x P(x)$ exists and which keeps to classical logic.[3] On such a theory, there can be no such property as $\lambda x(x \text{ is a red ball} \vee x \text{ is a property that doesn't instantiate itself})$. For properties aren't red balls (at least in the actual world, and let's focus on that), so by (ii), if such a property existed it would instantiate itself if and only if it doesn't; and that can't happen in classical logic. Unfortunately, the non-existence of such a property is awkward. For that together with the (EXP) from which we started implies that 'is either a red ball or a property that doesn't instantiate itself' isn't true of red balls. Presumably that isn't acceptable, so presumably the non-existence theorist will reject (EXP).[4] I don't say that's devastating, but I do think

[3] Taking this to include all the usual structural rules, including the transitivity of inference.

[4] The exceptions to (EXP) would be far more pervasive on a non-existence solution to the property-theoretic paradoxes that models its view of property-existence on an iterative theory of classes. For on such a theory there can be no class that contains everything (including all classes); so if properties are conceived similarly, there can be no universal property (property instantiated by everything, whether or not a property). So there can be no property $\lambda x(x = x)$. Also, there can be no property $\lambda x \neg(x \text{ is an electron})$. Since presumably both '$x = x$' and '$\neg(x \text{ is an electron})$' are true of protons, we'd have exceptions to (EXP) even for non-paradoxical predicates. (Quine's *New*

that (EXP) is part of our usual conception of "what properties are good for", and a classical theorist is probably best off keeping (i) while restricting (ii) since this leaves (EXP) unthreatened. (A naive theory by definition keeps (i) as well as (ii), so on it there is no threat to (EXP).)

As I hinted, I will want to strengthen the requirements on a naive theory of properties, and correspondingly, a naive theory of propositions: I want to allow for properties and propositions that aren't straightforwardly definable in our language, but only definable from parameters. That is, for properties I will demand that for each formula $P(x; u_1, ..., u_n)$ of our language (with no ambiguous or indexical elements and) with $x$ free and allowing for $n$ additional free variables $u_1, ..., u_n$ for any $n \geq 0$, then for any entities $o_1, ..., o_n$,

**($i^+$):** there is a corresponding property $\lambda x P(x; o_1, ..., o_n)$, and

**($ii^+$):** $\lambda x P(x; o_1, ..., o_n)$ is instantiated (at a world $w$) by all and only those entities $o$ such that (at $w$) $P(o; o_1, ..., o_n)$.

(The analogous expansion ($i_0^+$) and ($ii_0^+$) is to be made in the naive theory of propositions.) I stress that there are no restrictions on the "parameters" $o_1, ..., o_n$: in particular, they might themselves be properties or propositions.

One advantage of allowing for parameters—and in particular, parameters that might themselves be properties or propositions—is that it immediately gives us compositional laws. Suppose we want the law

**(C):** For any properties $P$ and $Q$ there is a property $P$ *and* $Q$ that is instantiated (at any world) by all and only the things that (at that world) instantiate them both.

Before the expansion to include parameters, (i) and (ii) would have allowed us to derive specific instances of (C), such as

> There is a property of *being red and a ball* that is instantiated (at any $w$) by all and only the things that (at $w$) instantiate both the property of *being red* and the property of *being a ball.*

But that is insufficient for two reasons: (a) it isn't a general law, just a collection of instances; and (b) it doesn't even cover all the relevant instances, e.g. it doesn't cover any properties that don't happen to be expressable in our language. But with the expansion, everything is fine: apply ($i^+$) and ($ii^+$) to the predicate $R(x; u_1, u_2)$ "$x$ instantiates $u_1$ and $x$ instantiates $u_2$", and then restrict the universal generalizations over $o_1$ and $o_2$ in ($i^+$) and ($ii^+$) to properties, including ones not definable in our language even from parameters (if there are such properties).

And in fact there may not even be any properties that aren't definable in our language *from parameters*, when those parameters can include properties. That's because ($i^+$) implies that for each property $P$, there is a property $P^* =_{df} \lambda x(x$ instantiates $P)$; and by ($ii^+$), it is instantiated at each world by precisely those things that instantiate $P$. Provided we take necessary co-instantiation as sufficient for property-identity,[5] $P^*$ just is $P$, and so $P$ is in a trivial sense definable in the language from itself as parameter. This is obviously only of limited interest: if there

*Foundations*, though generally regarded as an unattractive set theory, would serve considerably better than iterative set theory as a model for a property theory on which to base semantics; but the problems in the main text hold for it too.)

[5]In the end, I won't quite do this. However, when $P$ is a property that can't be instantiated by properties or propositions, the proposals in Section 6 will take $P^*$ to just be $P$.

is a property of being ferschlugginer that is otherwise undefinable in our language, this gives no way of introducing it into our language in a way that lets us understand it. Still, it does allow the theory to apply to such properties even though we don't understand them.

It will simplify many of the formulations below to introduce the ideas of a *parameterized 1-formula* and a *parameterized sentence.* The latter is simply a pair of a formula and an assignment of objects (in a broad sense that includes properties and propositions) to *all* its free variables; the former is a pair of a formula and an assignment of objects to *all but one of* its free variables. Then $(i^+)$ and $(ii^+)$ are subsumed under (i) and (ii) provided that we understand $P(x)$ to be a *parameterized* 1-formula; and analogously for $(i_0^+)$ and $(ii_0^+)$ and parameterized sentences. For future use, I'll also let a *fully parameterized abstraction term* be a pair of an abstraction term together with an assignment of objects to all its free variables.

How fine-grained are properties and propositions to be? That is, what is the relation that $P(x)$ must bear to $Q(x)$ (where $P$ and $Q$ may contain parameters) for $\lambda x P(x)$ to equal $\lambda x Q x$, and what is the relation that $B$ must bear to $C$ for $\lambda B$ to equal $\lambda C$? I'm inclined to think that there is no one "right answer" to this question, that it is basically a matter for stipulation (though with some limitations on which stipulations are coherent).

That said, there is much to be said for developing a theory that is quite coarse grained, something very much in the spirit of the worlds account with which I started. It can't be exactly that; and I won't be able to deal with the matter until Section 6, when the basic theory is in place. Until then, I will simply leave the notion of property-identity out of the theory.

Two final remarks before we get going:

First, the theory to be offered here extends also to relations (viewed non-extensionally); indeed, relations might be thought of as multi-place properties and propositions as 0-place properties (whereas properties in the sense originally intended were 1-place). If we do so, it is possible to develop the naive theories of all three in a common way. But there is a slight awkwardness about the format of the common theory: if one wants a theory that deals with relations with arbitrarily many places, the natural way to do it is by an instantiation predicate with arbitrarily many places; but predicates of varying adicity present certain complications. The awkwardness certainly isn't insurmountable, but for simplicity I will leave relations out of my presentation of the basic theory. This is not much of a loss, since if we have arbitrary $k$-tuples available, we can slightly artificially think of a $k$-place relation as a property of $k$-tuples.

Second, though the official focus of this paper is on the non-linguistic (properties and propositions), the main problems that must be overcome in achieving naivety are the same in the linguistic case as in the non-linguistic. (The one exception is that the problems about property-identity don't arise in the linguistic case.) The reader who wishes can easily modify the discussion in Sections 2-5 to avoid all talk of properties and propositions and to instead focus on language.

## 2. Generalizing Kripke's Construction

I'm going to begin the formal discussion with a generalization of a familiar tool, from Kripke 1975. (It's more familiar in the context of sentential truth than of property-instantiation or propositional truth, but it's well-known that it can be applied to those as well.) The main generalization is that Kripke developed the tool in the context where the valuation space is the Kleene algebra: the set $\{0, \frac{1}{2}, 1\}$, ordered in the usual way, with the obvious involution operation (taking value $v$ to $1 - v$) to handle negation.[6] Visser 1984 generalized the value space for the construction in one way, to the context of the Dunn 4-algebra: the set {0, b, c, 1}, partially ordered in such a way that $\forall v(0 \leq v \leq 1)$ and that b is incomparable to c, with the involution operation that flips 1 with 0 and that takes each of b and c to itself. I'm going to generalize the algebra in a different way (which could be further generalized to incorporate Visser's, though I won't bother to do so). This generalization will turn out to be very useful when it comes to conditionals.

Let a *Kripke algebra* be a complete deMorgan algebra in which (unlike with the Dunn algebra) there is an element $\frac{1}{2}$ that is a fixed point of negation and is comparable to every element. More fully: it has the form $\langle V, \leq, 1, \frac{1}{2}, 0, \natural\rangle$ where[7]

**(i):** $V$ is a set with at least three distinct elements 0, $\frac{1}{2}$ and 1
**(ii):** $\leq$ is a partial ordering on $V$, with $0 \leq \frac{1}{2} \leq 1$
**(iii):** every subset of $V$ has a least upper bound in $V$ and a greatest lower bound in $V$
**(iv):** $(\forall v \in V)(0 \leq v \leq 1)$
**(v):** $\natural$ is an operation on $V$ such that$(\forall u, v \in V)$ $(\natural u \leq \natural v$ iff $v \leq u)$
**(vi):** $(\forall v \in V)(\natural\natural v = v)$
**(vii):** $\natural(\frac{1}{2}) = \frac{1}{2}$
**(viii):** $(\forall v \in V)(v \leq \frac{1}{2} \vee \frac{1}{2} \leq v)$

(The last condition together with the partial ordering entails that $\frac{1}{2}$ is the only fixed point of the involution (or "negation") $\natural$.)

One example of a Kripke algebra that will be useful to bear in mind (though special in that the order is total) has $V$ the unit interval [0,1], $\leq$ the usual order on it, and $\natural = 1 - v$. This is the example that I will put to use for conditionals, starting in Section 3.

### 2.1. Background Framework.

Let $L_0$ be a first order modal language (for simplicity I'll assume that it has no primitive function symbols), and $L_0^+$ the result of expanding it in the obvious way to include (i) an abstraction operator $\lambda$ for forming terms for properties and propositions, together with (ii) a 1-place predicate 'Property', another 1-place predicate 'Proposition', and a 2-place predicate $\xi$. (This is short of the full $L$ we'll later use, which includes also two binary operators '$\rightarrow$' and '$\triangleright$' on formulas.) For any formula $A$, $\lambda A$ is to be a singular term whose free variables are just those of $A$; relative to any assignment of objects to all the variables of $A$, it denotes a proposition. For any formula $A$ and any variable $x$, $\lambda x A$ is to be a singular term whose free variables are just those *other than* $x$ that are free in $A$; relative to any assignment of objects to those other variables, it denotes a property. (We could if we like restrict to the case where $x$ is free in $A$, it won't matter. But if we don't, then even if $x$ isn't free in $A$, I'll regard $\lambda x A$ as denoting a property

[6]Kripke also considered supervaluationist alternatives, where the valuation space is a Boolean algebra, but this isn't suited to naive theories.

[7]There's a bit of redundancy here, which I've kept for the sake of clarity.

rather than a proposition; for instance, $\lambda x \forall y(Number(y))$ will denote the property of being in a world where everything is a number.)[8]

What does '$\xi$' mean? The parameterized formula '$o_1 \, \xi \, o_2$' is to be regarded as false whenever $o_2$ isn't a proposition or property. When $o_2$ is a property, '$o_1 \, \xi \, o_2$' can be read as '$o_1$ instantiates $o_2$'. When $o_2$ is a proposition, '$o_1 \, \xi \, o_2$' can be read as '$o_2$ is true'; in this case, $o_1$ won't matter. So '$\xi$' embodies truth and instantiation together: $True(y)$ is to be an abbreviation of "$y$ is a proposition and $\forall x(x \, \xi \, y)$" (though we could replace the $\forall$ with an $\exists$ without affecting anything); and $Instantiates(x, y)$ as an abbreviation of "$y$ is a property and $x \, \xi \, y$".

Let $M_0$ be any $V$-valued modal model for $L_0$. More precisely, $M_0$ will consist of

**(i):** a non-empty set $W$ of worlds
**(ii):** for each $w \in W$, a subset $W_w$ of $W$ (the set of worlds "accessible from $w$")
**(iii):** for each $w \in W$, a set $|M_0|_w$ of objects (the domain of $w$). I'll let $|M_0|$ be $\bigcup_{w \in W} |M_0|_w$, and refer to this as the domain of the modal model.
**(iv):** for each name of $L_0$, a member of $|M_0|$
**(v):** for each $k$-place predicate $p$ of $L_0$, and each $w \in W$, a function $p_w$ from $|M_0|^k$ to $V$. ("The $V$-valued extension of $p$ at $w$, in $M_0$".)

(For simplicity, I'm allowing worlds to overlap, so that one doesn't need a counterpart relation for cross-world identifications. If one likes one could impose "actualist" restrictions, such as that if neither $b$ nor $c$ is in $|M_0|_w$ then $p_w(o_1, ..., o_{i-1}, b, o_{i+1}, ..., o_k) = p_w(o_1, ..., o_{i-1}, c, o_{i+1}, ..., o_k)$; but I won't bother.) For the definition of validity (see Section 2.3), I also include

**(vi):** a nonempty subset $N$ of $W$: the set of "normal" worlds.

But $N$ could just be $W$. For simplicity in the construction to follow, I'll assume that nothing in $|M_0|$ is a term or formula of $L_0^+$, or a set built in part out of such terms or formulas. There will be no loss of generality, since if we have a modal model that doesn't meet this condition we can replace it by an isomorphic one that does.

We want a modal model $M_0^+$ (which I'll often just call $M$) with a domain that extends $|M_0|$ to include representatives of properties and propositions, but which looks like $M_0$ on the domain of the latter. (I emphasize that this is only a model: it makes no claim to include representatives of all properties in its domain.) For simplicity I'll do the construction in such a way that any proposition or property in the domain of the expanded model is in the domain of every world of the model; so that even if $o$ isn't in a given world, a property defined using $o$ as a parameter is. This won't play any essential role in the theory, it just makes for simplicity.

The expansion of the domain of $M_0$ goes in successive stages: in each stage, we construct representatives of properties and propositions from formulas, allowing parameters both from the ground model and from prior stages. (The stratification into $n$-formulas isn't used; this is a different stratification.) More fully: let

[8]The formation rules are given in cumulative levels. The *0-terms* are just the names and variables of $L_0$. From these we use the usual formation rules for the connectives to construct *0-formulas*, which will be just the formulas with no abstraction terms. The *1-terms* will be the 0-terms together with abstraction terms formed from 0-formulas. (They can have free variables, to be filled by parameters.) From these we construct the *1-formulas*, and then *2-terms*, which are the 0-terms together with abstraction terms formed from 1-formulas. And so on. The terms and formulas are whatever appears at some finite level.

$PROPOS[X]$= {parameterized sentences whose parameters are in $X$}, and

$PROPTY[X]$= {parameterized 1-formulas whose parameters are in $X$}.

For any natural number $j$, let

$PROPOS_j = PROPOS[|M_0| \cup \bigcup\{PROPOS_k \cup PROPTY_k : k < j\}]$, and

$PROPTY_j = PROPTY[|M_0| \cup \bigcup\{PROPOS_k \cup PROPTY_k : k < j\}]$.

Let $PROPOS_\omega$ be the union of the $PROPOS_j$, and similarly for $PROPTY_\omega$. (These depend on the starting model $M_0$: we shouldn't expect the model to contain a representation of a property of being ferschlugginer, if that is intuitively an "object level" property and there is no corresponding formula in the ground language. But this is no problem, we're only constructing a model, not making a claim about reality.) The members of $PROPOS_\omega$ and $PROPTY_\omega$ are *representations of* propositions and properties: we should think of the representation of properties and propositions as many-one, but we're not yet in a position to specify the equivalence relation on representations that make them representations of the same property or proposition. In Section 6 I'll consider ways of defining property identity in the language, with a formula $R$ whose extension in the model is the desired equivalence relation.

Until further notice, the domain of $M_0^+$ is to be $|M_0| \cup PROPOS_\omega \cup PROPTY_\omega$. Once a definition of property identity is in place, we can contract the model. And (following the simplifying stipulation of two paragraphs back) each $|M_0^+|_w$ is also just $|M_0|_w \cup PROPOS_\omega \cup PROPTY_\omega$.

The next step in the specification of $M_0^+$ is the treatment of denotation. I'll assign denotation to fully parameterized terms of $L_0^+$. Since the only terms are $L_0$-names, abstraction terms, and variables, these are just the $L_0$-names plus parameterized closed abstraction terms plus "parameterized variables". The latter are in effect proper names; so we in effect have a name for every object in the domain $|M_0^+|$. In $M_0^+$, each name of $L_0$ will denote what it denotes in $M_0$; and each parameterized closed abstraction term will denote a member of $PROPOS_\omega$ or $PROPTY_\omega$.

To complete the specification of the modal model $M$, we need to assign a value in $V$ to every parameterized sentence at each world. That's where Kripke comes in.

## 2.2. The Kripke Construction Generalized to Kripke Algebras.

As generalized to arbitrary Kripke algebras, Kripke's construction for the modal language under consideration is as follows. As a preliminary:

Let an *OPW triple* ("Object, Property/proposition, World", where again "objects" include properties and propositions) be a member of $|M_0^+| \times (|M_0^+| - |M_0|) \times W$.

Let a *valuation for* $\xi$ be any function $I$ that assigns a member of $V$ to each OPW triple.

*Relative to such an* $I$, we evaluate parameterized sentences at worlds by generalized Kleene rules:

- For any $k$-place predicate $p$ of $L_0$ and fully parameterized terms $t_1, ..., t_k$: (i) if for each $i$, the denotation $o_i$ of $t_i$ is in the ground model $|M_0|$, then

$|p(t_1, ..., t_k)|_{I,w}$ is the same as in $M_0$ (i.e. it's the value that $p_w$ assigns to $\langle o_1, ..., o_k\rangle$);
(ii) if for at least one $i$ the denotation $o_i$ of $t_i$ is a property or proposition, then for each $w$, $|p(t_1, ..., t_k)|_{I,w} = 0$.

- $|Property(t)|_{I,w}$ is 1 if the object denoted by $t$ is (a representative of) a property; 0 otherwise. Analogously for $|Proposition(t)|_{I,w}$.
- *If* $t_2$ denotes a member of $|M_0|$ (i.e. isn't a property or proposition), then $|t_1 \,\xi\, t_2|_{I,w}$ is 0
- Otherwise $|t_1 \,\xi\, t_2|_{I,w}$ is $I(o_1, o_2, w)$, where $o_1$ and $o_2$ are the objects denoted by parameterized terms $t_1$ and $t_2$.
- $|\neg A|_{I,w} = \natural(|A|_{I,w})$
- $|A \wedge B|_{I,w} = glb\{|A|_{I,w}, |B|_{I,w}\}$
- $|A \vee B|_{I,w} = lub\{|A|_{I,w}, |B|_{I,w}\}$
- $|\forall x A|_{I,w} = glb\{|A(o/x)|_{I,w} : o\ in\ |M_0^+|\}$, where $A(o/x)$ is the parameterized formula just like $A$ except with $o$ assigned to the free variable $x$
- $|\exists x A|_{I,w} = lub\{|A(o/x)|_{I,w} : o\ in\ |M_0^+|\}$
- $|\Box A|_{I,w} = glb\{|A|_{I,w^*} : w^* \in W_w\}$
- $|\Diamond A|_{I,w} = lub\{|A|_{I,w^*} : w^* \in W_w\}$.

(The clauses for quantifiers and modal operators depend on the completeness assumption (iii) for Kripke algebras.)

Note that if $L_0$ contains an identity predicate, the rule for atomic formulas dictates that in $L_0^+$ it is not treated as a full identity predicate but as identity restricted to the ground model. Identity among properties and propositions is not yet defined.

Now for the crucial definition, whose interest depends on clauses (vii) and (viii) in the definition of Kripke algebras. If $I$, $J$ are valuations for $\xi$, let $I \leq_K J$ mean:

For all OPW triples $\langle o_1, o_2, w\rangle$, either $\frac{1}{2} \leq I(o_1, o_2, w) \leq J(o_1, o_2, w)$, or else $J(o_1, o_2, w) \leq I(o_1, o_2, w) \leq \frac{1}{2}$.

The importance of this is that it allows us to generalize Kripke's key Lemma to arbitrary Kripke algebras:

**Kripke Monotonicity Lemma (Generalized Form):** if $I$ and $J$ are valuations for $\xi$ with $I \leq_K J$, then for every parameterized sentence $B$ and world $w$, either $\frac{1}{2} \leq |B|_{I,w} \leq |B|_{J,w}$ or else $|B|_{J,w} \leq |B|_{I,w} \leq \frac{1}{2}$.

**Proof:** By induction on complexity of $B$.

- When $B$ atomic, then the only case where the value depends on $I$ is when $B$ has form $t_1 \,\xi\, t_2$ for parameterized terms, and $t_2$ denotes a property or proposition. And in that case, $|B|_{I,w}$ is just $I(den(t_1), den(t_2), w)$. Similarly for $J$. And so the assumption that $I \leq_K J$ yields the desired conclusion that either $\frac{1}{2} \leq |B|_{I,w} \leq |B|_{J,w}$ or else $|B|_{J,w} \leq |B|_{I,w} \leq \frac{1}{2}$.
- Negation: If $\frac{1}{2} \leq |B|_{I,w} \leq |B|_{J,w}$ then by order-reversing nature of $\natural$, $(|B|_{J,w})^{\#} \leq (|B|_{I,w})^{\#} \leq \frac{1}{2}$, i.e. $|\neg B|_{J,w} \leq |\neg B|_{I,w} \leq \frac{1}{2}$. Similarly for the other case.
- Conjunction: Suppose $B$ and $C$ both satisfy the claim.
  Case 1: $\frac{1}{2} \leq |B|_{I,w} \leq |B|_{J,w}$ and $\frac{1}{2} \leq |C|_{I,w} \leq |C|_{J,w}$. Then $\frac{1}{2} \leq |B \wedge C|_{I,w} \leq |B \wedge C|_{J,w}$.
  Case 2: $|B|_{J,w} \leq |B|_{I,w} \leq \frac{1}{2}$ and $|C|_{J,w} \leq |C|_{I,w} \leq \frac{1}{2}$. Then $|B \wedge C|_{J,w} \leq |B \wedge C|_{I,w} \leq \frac{1}{2}$.

Case 3: $|B|_{J,w} \leq |B|_{I,w} \leq \frac{1}{2}$ and $\frac{1}{2} \leq |C|_{I,w} \leq |C|_{J,w}$. Then $|B \wedge C|_{J,w} = |B|_{J,w}$ and $|B \wedge C|_{I,w} = |B|_{I,w}$, so result holds.
The other case is analogous to Case 3.

- Universal quantification and $\Box$ are similar. For instance, if $B(o)$ obeys the assumption for all $o$, then
(i) if all the $|B(o)|_{I,w}$ are at least $\frac{1}{2}$ then for all $o$ $\frac{1}{2} \leq |B(o)|_{I,w} \leq |B(o)|_{J,w}$ and hence $\frac{1}{2} \leq |\forall x Bx|_{I,w} \leq |\forall x Bx|_{J,w}$;
(ii) otherwise, $|\forall x Bx|_{I,w}$ and $|\forall x Bx|_{J,w}$ are the glbs of the $|B(o)|_{I,w}$ and $|B(o)|_{J,w}$ for which $|B(o)|_{I,w} < \frac{1}{2}$ . For these, $|B(o)|_{J,w} \leq |B(o)|_{I,w} \leq \frac{1}{2}$, and so $\frac{1}{2} \leq |\forall x Bx|_{J,w} \leq |\forall x Bx|_{I,w}$.
- Disjunction and existential quantification and $\Diamond$ are similar. ■

Given any valuation $I$ for $\xi$, we define its "Kripke jump" $K(I)$ as follows:

> $K(I)$ is the valuation for $\xi$ that, for each $o$ in the full $|M_0^+|$ and each parameterized sentence or parameterized 1-formula $A$ and each world $w$, assigns to $\langle o, \lambda z A(z), w\rangle$ or $\langle o, \lambda A, w\rangle$ the value $|A(o)|_{I,w}$.

(Obviously $K(I)$ makes all propositions "instantiation invariant": if $o\,\xi\, c$ where $c$ represents a proposition, then for any other $o^*$, $o^*\,\xi\, c$.)

The following is then just a restatement of the Monotonicity Lemma:

**Kripke Monotonicity Corollary:** if $I$ and $J$ are valuations for $\xi$ with $I \leq_K J$, then $K(I) \leq_K K(J)$.

Moreover, the completeness requirement (iii) on Kripke algebras allows us to define a valuation for $\xi$ corresponding to any sequence of valuations for $\xi$, say as follows:

> $LIM\{I_\rho : \rho < \tau\}$ is the valuation for $\xi$ that assigns to any OPW triple $\langle o_1, o_2, w\rangle$ the following:
> $lub\{I_\rho(o_1, o_2, w) : \rho < \tau\}$, if $(\exists \rho < \tau\ )(I_\rho(o_1, o_2, w) > \frac{1}{2})$
> $glb\{I_\rho(o_1, o_2, w) : \rho < \tau\}$, if $(\forall \rho < \tau)(I_\rho(o_1, o_2, w) \leq \frac{1}{2})$.

This isn't of much interest in general, but it is when the sequence $\{I_\rho : \rho < \tau\}$ is an $\leq_K$*-chain* of valuations for $\xi$; that is, a sequence of valuations for $\xi$ such that for any $\rho$ and $\sigma$ such that $\rho < \sigma < \tau$, $I_\rho \leq_K I_\sigma$. In that case we have:

**Observation on Chains:** If $\{I_\rho : \rho < \tau\}$ is an $\leq_K$-chain, then for any $\sigma < \tau$, $I_\sigma \leq_K LIM\{I_\rho : \rho < \tau\}$.

**Proof:** Suppose that $\{I_\rho : \rho < \tau\}$ is an $\leq_K$-chain and $\sigma < \tau$; and consider any $\langle o_1, o_2, w\rangle$. We need (i) that if $I_\sigma(\langle o_1, o_2, w\rangle) > \frac{1}{2}$ then $LIM\{I_\rho : \rho < \tau\})(\langle o_1, o_2, w\rangle) \geq I_\sigma(\langle o_1, o_2, w\rangle)$, and (ii) that if $I_\sigma(\langle o_1, o_2, w\rangle) < \frac{1}{2}$ then $LIM\{I_\rho : \rho < \tau\})(\langle o_1, o_2, w\rangle) \leq I_\sigma(\langle o_1, o_2, w\rangle)$. (i) is immediate from the definition of LIM, independent of the chain assumption. (ii) holds because $I_\sigma(\langle o_1, o_2, w\rangle) < \frac{1}{2}$ together with the chain condition implies that for all $\rho < \tau$, $I_\rho(\langle o_1, o_2, w\rangle) \leq \frac{1}{2}$; and so $LIM\{I_\rho : \rho < \tau\})(\langle o_1, o_2, w\rangle)$ is the glb of the $I_\rho(\langle o_1, o_2, w\rangle)$ and hence $\leq I_\sigma(\langle o_1, o_2, w\rangle)$. (Of course the asymmetry in the proof is just due to the asymmetry in the definition, an asymmetry that only makes a difference in the uninteresting case of non-chains.) ■

We can now easily prove the existence of Fixed Points, and in particular a minimal one. For the latter, let $I_0$ be the trivial valuation for $\xi$ that assigns $\frac{1}{2}$ to every triple in its domain; whenever $I_\rho$ has been defined, define $I_{\rho+1}$ as $K(I_\rho)$; and when $I_\rho$ has been defined for $\rho < \lambda$ when $\lambda$ is a limit ordinal, let $I_\lambda$ be $LIM\{I_\rho : \rho < \lambda\}$. By the Monotonicity Corollary and the Observation on Chains, it is clear that whenever $\rho < \sigma$, $I_\rho \leq_K I_\sigma$. So there must be an ordinal $\tau$ of cardinality no greater than that of the set of all valuations for $\xi$ such that $I_{\tau+1} = I_\tau$. This is the minimal fixed point. (If at stage 0 we had started out not with the trivial valuation, but with another $I^*$ such that $I^* \leq_K K(I^*)$, we would reach a fixed point by the same argument, the minimal valuation *that extends* $I^*$.) Restating (in a slightly loose notation using "parameterized formulas" to avoid talk of functions assigning objects to variables):

**Kripke Fixed Point Theorem (Generalized to Arbitrary Kripke Algebras):** There are valuations $I$ for $\xi$ (including one that is minimal in the ordering $\leq_K$) such that for each $o, o_1, ..., o_k$ in the full $|M_0^+|$ and each formula $A$ that has exactly $k$ variables other than $x$ free (and may or may not have $x$ free),
$|o\,\xi\,\lambda x A(x, o_1, ..., o_k)|_{I,w} = |A(o, o_1, ..., o_k))|_{I,w}$,
and for each without $x$ free,
$|o\,\xi\,\lambda A(o_1, ..., o_k)|_{I,w} = |A(o_1, ..., o_k))|_{I,w}$.

(The left hand side is what $I_\tau$ assigns to $\langle o, \lambda x A(x, o_1, ..., o_k), w\rangle$, the right hand side is what $I_{\tau+1}$ assigns it, and $I_{\tau+1} = I_\tau$.)

By using such fixed points, then, we are guaranteed $(ii^+)$ and $(ii_0^+)$ of the naive theory of properties and propositions. (Recall the definitions of 'instantiates' and 'True' in terms of $\xi$.) $(i^+)$ and $(i_0^+)$ were built into the construction from the start.

This generalization of the Kripke construction on $\{0, \frac{1}{2}, 1\}$ to arbitrary Kripke algebras isn't particularly surprising, but it is useful.[9] A minor illustration (a more substantial one will be given in Section 3): Suppose we think (as many people do) that vague language is best evaluated with values in the unit interval [0,1] (and not confined to 0, $\frac{1}{2}$ and 1), with the rules for conjunction, disjunction and the quantifiers given in terms of greatest lower bound and least upper bound and where $|\neg A|$ is $1 - |A|$. Then it's natural to hope that we can add propositions and properties to a language with vague terms in a naive way, without disrupting the values of the ground level sentences. Since [0,1] with this negation is a Kripke algebra, the generalized Kripke construction gives just what we need.

2.3. **Choices for Validity in the Conditional-free Context.** The construction I've been developing can be used as a model-theoretic semantics for many different logics, for there are many different choices for how to use it to define validity. In the ones I'll be interested in, validity is defined in terms of the values of parameterized formulas *at all normal worlds of all models.*

One kind of choice is what constraints if any one puts on the modal structure. I'll eventually impose the reflexivity condition for normal worlds: that at least for normal worlds, $w \in W_w$. That's what's needed for the validity of $\Box A \models A$. Also, to simplify a discussion later on I will assume that if there are non-normal worlds then for each of them, there is a normal world from which it is accessible. The point of this is to make the claim $\models \Box B$ require that $B$ is true at *every* world of every model, even the non-normal ones.

But a more fundamental choice, with the kind of multi-valued semantics considered here, is how exactly the values enter into the account of validity. This greatly affects the logic that the semantics validates. For instance, if we take an inference

[9]Further generalizations are possible: e.g. for any Kripke algebra $V$, we can add a value incomparable to all the values in $V$ other than 1 or 0, and let the negation involution $\natural$ take it to itself; this generalizes the way Visser went from the Kleene 3-algebra to the Dunn 4-algebra. Also, the Fixed Point Theorem as stated above automatically extends to products and subproducts of Kripke algebras (which typically will not themselves be Kripke algebras due to condition (viii)).

to be valid if in all models and all normal worlds in them, if the premises have value 1, so does the conclusion, then we'll get a "paracomplete" logic in which excluded middle isn't valid. If we take an inference to be valid if in all models and normal worlds in them where the conclusion has value 0, so does at least one of the premises, then we'll get a "paraconsistent" logic in which disjunctive syllogism (the inference from $A \vee B$ and $\neg A$ to $B$) isn't valid. (Similarly if we replace '0' by 'less than $\frac{1}{2}$'.) If we take an inference to be valid if in all models and normal worlds in them, the value of the conclusion is at least the greatest lower bound of the values of the premises, we have both failures: the logic is both paracomplete and paraconsistent. Nothing said so far is a reason for going one way or another. And indeed, that is an issue on which I will mostly remain neutral in this paper, though there will be some remarks in Section 5 tending to favor the paracomplete approach.

## 3. Restricted Quantifier Conditionals: Revising the Revision Approach.

Much of the recent work on naive theories of truth and property-instantiation has been devoted to conditionals. As first noted in Beall *et al* 2006 and Chapter 18 of Priest 2006, an adequate theory needs to deal with at least two kinds of conditionals: one (which I'll symbolize as $\rightarrow$) for defining restricted universal quantification from unrestricted,[10] another (which I'll symbolize as $\triangleright$) to symbolize the kind of conditional we use in everyday sentences like

> He's under the delusion that if he runs for President, he'll win. But he doesn't want the job, so he won't run.

In a classical setting, $\triangleright$ is obviously not the material $\supset$, whereas the $\rightarrow$ is. In a non-classical setting of the sort required for naive properties and truth, even the $\rightarrow$ can't obey the usual classical definition of $\supset$ (that is, $\neg A \vee B$), if the $\rightarrow$ is to obey reasonable laws (modus ponens and $A \rightarrow A$). But presumably it needs to reduce to $\supset$ when the antecedent and consequent behave classically. Whereas $\triangleright$ had better *not* reduce to $\supset$ in such classical contexts.

There are various approaches to handling each of the two conditionals (and to handling how they interact). One approach that can be used for each is a revision construction. In the past I've developed this approach in the context of a valuation space based on $\{0, \frac{1}{2}, 1\}$ (resulting in a much bigger valuation space that is a product or subproduct of $\{0, \frac{1}{2}, 1\}$, that is, a space of functions from some set $X$ to $\{0, \frac{1}{2}, 1\}$). But it can be adapted for other Kleene algebras in place of $\{0, \frac{1}{2}, 1\}$, with results that have a rather different feel. In the rest of this section I'll use [0,1] rather than $\{0, \frac{1}{2}, 1\}$ in this role of "mini-space" (so that the resulting valuation space is a space of functions from some set $X$ to [0,1]).[11]

---

[10] Instead of defining "All $A$ are $B$" as $\forall x(Ax \rightarrow Bx)$ for some appropriate $\rightarrow$, we could take a binary restricted quantifier as primitive. But that wouldn't effect the logical issues to be discussed. For we could then use primitive restricted quantification to define a conditional $\rightarrow$: $A \rightarrow B$ would mean that all $v$ such that $A$ are such that $B$, where $v$ is a variable not free in either $A$ or $B$; and with this defined $\rightarrow$, $\forall x(Ax \rightarrow Bx)$ will be equivalent to the original "All $A$ are $B$", assuming only very uncontroversial laws.

[11] There are other approaches that have been used with the 3-valued mini-space. The earliest is due to Brady, and I will make use of an improved version of it in Section 4.3.2, in connection with the $\triangleright$ conditional. For reasons to be mentioned there, I don't think the Brady approach has

In this section I'll illustrate this for the restricted quantifier conditional $\rightarrow$, saving the $\triangleright$ for the next section.

There is an initial worry about how to make the adaptation from $\{0, \frac{1}{2}, 1\}$ to [0,1].

The basic idea of the revision construction (in a language that adds $\rightarrow$ to the $L_0^+$ of Section 2) is that we start out with an assignment $h$ of values in [0,1] to parameterized conditional sentences at worlds (that is, to pairs of conditional formulas and functions assigning objects to their free variables, at worlds). We require of $h$ that it be "transparent", in the sense that if $C_1$ and $C_2$ are conditional formulas such that one results from another by substitutions of formulas of form "$o\,\xi\, B(x, o_1, ..., o_k)$" for the corresponding "$B(o, o_1, ..., o_k)$" or vice versa, then for any world and any common assignment of objects to the variables of $C_1$ and $C_2$, $h$ assigns the same value to the parameterizations of $C_1$ and $C_2$ at that world. Given such a transparent valuation $h$ of (parameterized) conditionals, we first do a generalized Kripkean construction as in the previous section, but in the language with $\rightarrow$, holding the values of parameterized conditionals fixed at the values given by $h$ throughout the construction. We choose a fixed point: say the minimal fixed point. For any sentence $A$, let $|A|_{h,w}$ be the Kripke fixed point value that $A$ gets at world $w$ on the construction starting from $h$. It is clear that because $h$ is transparent, the Kripke construction guarantees that the assignment $|A|_{h,w}$ that it generates is transparent in the corresponding sense. The key is then to introduce a revision rule, that takes any $h$ to a new one $R(h)$, based on the values $|A|_{h,w}$, which will be transparent given that $h$ is. (The rule for $\rightarrow$ operates world by world: that is, the value that $R(h)$ assigns to a conditional at a given $w$ depends only on the values that $h$ assigns its antecedent and consequent at that same $w$. The situation will be different when we come to $\triangleright$.)

More specifically, the revision rule I've come to advocate for $\rightarrow$ in the case of the space $\{0, \frac{1}{2}, 1\}$ is this:[12]

much relevance to restricted quantification, and Brady doesn't either (judging from Beall *et al* 2006, of which he was a co-author). [But see note 33.]

Another alternative to revision semantics is the "higher order fixed point" approach of Field 2014: higher order in that the fixed point isn't an individual hypothesis but a set of final hypotheses together with their ordering. This approach can be used for either conditional. The Field 2014 version uses valuations in a space of form $\{0, \frac{1}{2}, 1\}^X$ though in this case $X$ will not be a set of ordinals; it too can be generalized to $[0, 1]^X$. This approach has deep structural similarities to the revision-theoretic approach. Indeed, as Harvey Lederman remarked to me, it selects the unique "minimal" fixed point in a higher order construction in which the revision constructions with different starting hypotheses produce the "maximal" fixed points. (Perhaps the terms 'maximal' and 'minimal' should be reversed: the set of hypotheses that survive in the revision-theoretic systems that I've called 'maximal' are proper subsets of those in the "minimal" fixed point. Nonetheless, the one I've called 'minimal' is the one generated most easily, and it produces the fewest sentences with extreme values 0 and 1.)

The higher order fixed point approach is more complicated than the revision, and especially once one shifts to the [0,1]-valued versions I'm not sure that the extra complications have much payoff. (In the $\{0, \frac{1}{2}, 1\}$ case, Standefer 2015 has pointed to some oddities in the revision approach that don't apply to the higher order fixed point approach; but the oddities are avoided in the [0,1]-valued revision approach as well.) I won't discuss the higher order approach further.

[12]In Field 2008 I primarily used a different 0 clause (though I mentioned the one in the text in Section 17.5): the 0 clause there (generalized to the current modal framework) was "0 iff $|A|_{h,w} > |B|_{h,w}$". The one in the text (also employed in Field 2014 and Field 2016) is substantially

$$[R(h)](A \to B, w) \text{ is } \begin{cases} 1 & \text{if } |A|_{h,w} \leq |B|_{h,w} \\ 0 & \text{if } |A|_{h,w} = 1 \text{ and } |B|_{h,w} = 0 \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

The natural generalization of this to [0,1] is

$$[R(h), w](A \to B) \text{ is } \begin{cases} 1 & \text{if } |A|_{h,w} \leq |B|_{h,w} \\ 1 - (|A|_{h,w} - |B|_{h,w}) & \text{otherwise.} \end{cases}$$

But the problem is that if, as is natural, we start the construction with a function $h_0$ that assigns to every conditional one of the values in $\{0,\ \frac{1}{2},\ 1\}$ at each world, then this rule for $R(h)$ will also assign only values in $\{0,\ \frac{1}{2},\ 1\}$; the extra richness in the space [0,1] won't be exploited.

The best way around this, I think, is to insist on "slow corrections" in the revision process. More fully, $R(h)$ as given above looks like it gives natural values based on the old $h$; but it produces big jumps, sometimes from 1 to 0 or 0 to 1, which we might have to reverse later. Rather than make the big jump at once, let's have a rule that averages the $R(h)$ value with the $h$ value:

$$[R^*(h), w](A \to B) \text{ is } \begin{cases} \frac{1}{2}[h(A \to B, w) + 1] & \text{if } |A|_{h,w} \leq |B|_{h,w} \\ \frac{1}{2}[h(A \to B, w) + 1 - (|A|_{h,w} - |B|_{h,w})] & \text{otherwise.} \end{cases}$$

That's the "slow correction" process.

Let us now use $R^*(h)$ to construct a revision sequence. Let $h_0$ be any transparent valuation: say the one that gives every conditional the value $\frac{1}{2}$ at every world, though I'll suggest what I regard as a better alternative in note 16. (For most purposes the details of $h_0$ won't matter much, as long as it is transparent: these details are very largely washed away as the revision construction proceeds, though there are a few special sentences for which it matters.) Once $h_\mu$ has been constructed, let $h_{\mu+1}$ be $R^*(h_\mu)$.

What about limits? It turns out that a great many $A$ and $B$ are such that for each world $w$, the sequence $\{|A \to B|_{h_n,w} : n < \omega\}$ approaches a particular point $r_w$ as limit; and in that case we presumably want to take that $r_w$ as the value that $h_\omega$ assigns the conditional $A \to B$ at $w$. More generally when there is convergence prior to any limit ordinal $\lambda$, that should determine the value at $\lambda$. But what about when there's no convergence? One might explore taking $h_\lambda(A \to B, w)$ to be the average of the liminf and limsup of $\{h_\mu(A \to B, w) : \mu < \lambda\}$. But I prefer a limit rule where it is to be the liminf when that is at least $\frac{1}{2}$, the limsup when that is at most $\frac{1}{2}$, and $\frac{1}{2}$ in all other cases, i.e. when the liminf is less than $\frac{1}{2}$ and the limsup more than $\frac{1}{2}$. ("The value is as close to $\frac{1}{2}$ as it can sensibly be.") Either of these rules generalizes the limit rule I've previously used in the case of $\{0,\ \frac{1}{2},\ 1\}$.[13]

---

better in several respects. In any case, the basic point next to be made doesn't depend on the difference.

(The 1 clause as given makes $\to$ contraposable, which I think reasonable. In the 3-valued case, there is also a natural non-contraposable $\to$ to consider, where the 1-clause is altered to: "1 iff either $|A|_{h,w} < 1$ or $|B|_{h,w} = 1$". But this doesn't generalize as neatly to the continuum-valued case.)

[13]A *prima facie* advantage of the averaging rule over the one I prefer is that it would lead outside the space $\{0,\ \frac{1}{2},\ 1\}$ even without "slow corrections" in the revision rule. But it doesn't lead outside $\{0,\ \frac{1}{2},\ 1\}$ as much as is desirable: later in this section I note that the resolution of "ordinary" paradoxes on the present semantics reduces to the Łukasiewicz resolution, and this attractive feature of the semantics depends on its use of slow corrections. (It doesn't depend on

Whichever of these limit rules we use, the general theory of revision sequences (Gupta and Belnap 1993) tells us that there are some "hypotheses" that occur arbitrarily late in the sequence; call these *recurring hypotheses.* (That is, $h_\kappa$ is recurring iff for any $\varsigma$, there is an $\eta > \varsigma$ for which $h_\eta = h_\kappa$.) Indeed, there are ordinals $\mu$ such that for any $\kappa \geq \mu$, $h_\kappa$ is recurring; call such $\mu$ *final.* And among these final ordinals, there are ones of particular interest, the *reflection ordinals*: these are the final limit ordinals $\Delta$ such that for any $\mu < \Delta$, and any final $\eta$, there is a $\kappa \in [\mu, \Delta)$ such that $h_\kappa = h_\eta$. And it's easy to see that for any two reflection ordinals, the values of all parameterized conditionals and hence of all parameterized sentences are the same.[14] More generally, on the preferred limit rule, it's easy to see that if $\Delta$ is a reflection ordinal and $\mu$ any other final ordinal, then:

**(*):** For any parameterized conditional $A \rightarrow B$ and world $w$, either $\frac{1}{2} \leq h_\Delta(A \rightarrow B, w) \leq h_\mu(A \rightarrow B, w)$ or $h_\mu(A \rightarrow B, w) \leq h_\Delta(A \rightarrow B, w) \leq \frac{1}{2}$.

And one can then, by an induction on complexity, extend this to the values of non-conditional parameterized sentences:

**(FT):** For any parameterized sentence $A$ and world $w$, either $\frac{1}{2} \leq |A|_{\Delta,w} \leq |A|_{\mu,w}$ or $|A|_{\mu,w} \leq |A|_{\Delta,w} \leq \frac{1}{2}$.

The proof is a straightforward generalization of the one I've given elsewhere (e.g. Field 2008) for the special case of $\{0, \frac{1}{2}, 1\}$: by induction on the stages $\sigma$ of the Kripke construction, with a subinduction on the complexity of parameterized sentences $A$, one proves that (deleting the world parameter for readability) if $|A|_{\Delta,\sigma} < \frac{1}{2}$ then for all final $\alpha$ $|A|_\alpha \leq |A|_{\Delta,\sigma}$, and if $|A|_{\Delta,\sigma} > \frac{1}{2}$ then for all final $\alpha$ $|A|_\alpha \geq |A|_{\Delta,\sigma}$.[15]

A special case of (*) is that if a conditional gets value 1 at a reflection ordinal, it gets value 1 for every final ordinal; and similarly for 0. And a special case of (FT) is that that's so for every parameterized sentence of the language. Thus (FT) is a generalized version of what in the 3-valued case I've called the "Fundamental Theorem". (These special cases of (*) and of (FT) would hold also on the alternative limit rule where we average the liminf and limsup, and it may be that only the special cases are crucial. But in what follows I'll use the preferred rule.)[16]

---

the choice between the two limit rules, but if we want the reduction, it removes the *prima facie* advantage of the averaging rule.)

[14]The term 'reflection ordinal' is sometimes used more broadly, for any ordinal $\kappa$ for which $h_\kappa$ is the same as $h_\Delta$ where $\Delta$ is as described here. The difference won't matter much, though confusion might result if it were not realized that there are infinitely many reflection ordinals in the latter sense between any two reflection ordinals in the former.

[15]There are only two places where the move to the space [0,1] might be thought to matter to the proof: for the clause for 'True' at limit ordinals in the main induction, and in the quantifier clause in the subinductions. But in neither case is there a problem.

(1) Suppose for instance that $|True(t)|_{\Delta,\lambda} > \frac{1}{2}$, where $t$ denotes $C$. Then by the Kripke rules, it must be that for each $r$ in the open interval $(\frac{1}{2}, r)$, there is a $\tau < \lambda$ such that $|C|_{\Delta,\tau}$ is at least $r$. So by induction hypothesis, for each final $\alpha$ and each such $r$, $|C|_\alpha \geq r$. That can only be so if $|C|_\alpha \geq |True(t)|_{\Delta,\lambda}$; which by transparency means that $|True(t)|_\alpha \geq |True(t)|_{\Delta,\lambda}$.

(2) Suppose for instance that $|\forall x Bx|_{\Delta,\sigma} < \frac{1}{2}$. (The case where $|\forall x Bx|_{\Delta,\sigma} > \frac{1}{2}$ is slightly easier.) Then the set $\Sigma = \{o : |B(o)|_{\Delta,\sigma} < \frac{1}{2}\} \neq \emptyset$, and $|\forall x Bx|_{\Delta,\sigma}$ is $glb\{|B(o)|_{\Delta,\sigma} : o \in \Sigma\}$. And $|\forall x Bx|_\alpha \leq glb\{|B(o)|_{\Delta,\sigma} : o \in \Sigma\}$. By subinduction hypothesis, $|B(o)|_\alpha \leq |B(o)|_{\Delta,\sigma}$ for all $o \in \Sigma$, so $|\forall x Bx|_\alpha \leq |\forall x Bx|_{\Delta,\sigma}$.

[16]This is all independent of the choice of transparent starting valuation $h_0$, and as I've said, that choice has only a minor effect on the results. But for the record, I now prefer to first do the construction in a general way, independent of the transparent starting valuations $h_0$. For

The upshot is that the value in $[0,1]$ of a parameterized sentence $A$ at a reflection ordinal $\Delta$ tells us a lot about how $A$ behaves in the model, but the full story requires how it behaves in a semi-open interval $[\Delta, \Delta+\Pi)$ between two reflection ordinals. (In such an interval, every recurring hypothesis shows up. It's best to think of it as the closed interval $[\Delta, \Delta+\Pi]$ but with endpoints identified to form a circle.) So the obvious value space is the space $[0,1]^{\Pi}$ of functions from $\Pi$ to $[0,1]$, where $\Pi$ is an ordinal that when added (on the right) to a reflection ordinal yields a bigger reflection ordinal. I'll let $||A||_w$ be the value in $[0,1]^{\Pi}$ of $A$ at $w$. For later reference, I'll use the notation $\mathbf{r}$ (where $r \in [0,1]$) for the constant function that assigns the value $r$ to every predecessor of $\Pi$. So for a sentence to have value $\mathbf{r}$ (at a world) is for it to have the same value $r$ (at that world) ***at all final ordinals***; it can have different values for earlier ordinals, but those earlier values get washed out.

How much of an improvement do we get in this construction by using [0,1] instead of $\{0, \frac{1}{2}, 1\}$ as the mini-space? In one sense, not much: the most general laws that don't hold with $\{0, \frac{1}{2}, 1\}$, such as the permutation axiom

$$[A \to (B \to C)] \to [B \to (A \to C)]$$

(or its weaker rule form), tend not to hold with [0,1] either. However, with [0,1] we have to go to much greater lengths to get exceptions: for "ordinary" paradoxical sentences, the new construction yields far more satisfactory results, and axioms such as permutation will hold for them.

That's because

> (i) for "ordinary" paradoxical sentences, their value in the new construction is the same for all $h_\mu$ where $\mu$ is final (often even, for all $h_\mu$ when $\mu \geq \omega$); that is, their value is one of the constant functions $\mathbf{r}$; and
>
> (ii) for such sentences, the semantics reduces to Łukasiewicz semantics on [0,1].

Łukasiewicz semantics on [0,1] uses the evaluation rules for $\neg$, $\wedge$, $\vee$ and the quantifiers and modal operators that were employed in Section 2, together with the rule

$$|A \to B|_{h,w} \text{ is } \begin{cases} 1 & \text{if } |A|_{h,w} \leq |B|_{h,w} \\ 1-(|A|_{h,w} - |B|_{h,w}) & \text{otherwise.} \end{cases}$$

The "jumpy correction" revision rule R was obviously modeled on this: it's essentially this except with $|A \to B|_{R(h),w}$ instead of $|A \to B|_{h,w}$ as its left hand side, so that what the jumpy revision rule semantics says of $R(h)$, Łukasiewicz semantics says of $h$ itself. The "slow correction" revision rule $R^*$ is still a further step from Łukasiewicz semantics. However, it's easy to see that if $A$ and $B$ are sentences each of whose values is a constant function (say $\mathbf{a}$ and $\mathbf{b}$), then neither step makes a difference: the value of $A \to B$ will itself be the constant function $\mathbf{c}$, where the value $c$ is determined from the values $a$ and $b$ by the Łukasiewicz rule.

---

each one, and each parameterized conditional $A \to B$ and world $w$, we get a reflection value $|A \to B|_{\Delta,w,h_0}$; let the set of these for a given $A \to B$ and $w$ but varying the starting hypothesis be $REFL(A \to B,w)$. Then my preferred starting valuation assigns to $A \to B$ at $w$ the greatest lower bound of $REFL(A \to B,w)$ if that's at least $\frac{1}{2}$, the least upper bound if that's no more than $\frac{1}{2}$, and $\frac{1}{2}$ in other cases. This choice gives more natural values to some sentences: e.g. to "conditional truth-teller" sentences whose antecedent is itself conditional, as on p. 319 of Yablo 2003. But its overall impact on the theory is small.

Łukasiewicz semantics doesn't include a specification for the generalized instantiation predicate $\xi$ (or for a truth predicate). But it's well-known that for the quantifer-free sublanguage (supplemented with a means to achieve self-reference and referential loops), a predicate that behaves naively in this sublanguage can be added.[17] Even outside this sublanguage, a great many paradoxical sentences can be consistently evaluated in Łukasiewicz semantics (often in a unique way). And it's natural to conjecture that for those sentences, the revision process given above will lead to the constant function corresponding to one of those values.[18] I won't attempt here to make this conjecture precise or to prove it, but there are countless examples to illustrate it. For instance, consider a Curry-like sentence $K_2$ that is equivalent to $True(\lambda K_2) \rightarrow [True(\lambda K_2) \rightarrow \bot]$.[19] Applying the Łukasiewicz rules without assuming naive truth, we get

$$|K_2|_{h,w} = \begin{cases} 1 & \text{if } |True(\lambda K_2)|_{h,w} \leq \frac{1}{2} \\ 2(1 - |True(\lambda K_2)|_{h,w}) & \text{if } |True(\lambda K_2)|_{h,w} > \frac{1}{2}. \end{cases}$$

We see that in Łukasiewicz semantics, at each world there is a unique value of $K_2$ where the naivety requirement that $|True(\lambda K_2)|_w = |K_2|_w$ is met, and it is the same at each world: it is $\frac{2}{3}$. (There is no world-dependence since $K_2$ is a "non-contingent" Curry-like sentence.) And in the modified revision semantics, it isn't hard to show (independent of what one assumes the starting valuation $h_0$ to be) that for all $\mu \geq \omega$, $h_\mu$ assigns value $\frac{2}{3}$ to $K_2$ at each world.[20] That value of course wasn't available on the revision construction using only $\{0, \frac{1}{2}, 1\}$. (There, the values repeat indefinitely in groups of three, each group consisting of $\frac{1}{2}$ followed by 1 followed by $\frac{1}{2}$.)[21]

Well, you might say, why not just use the Łukasiewicz semantics, it's simpler? The answer is basically that there's no way to expand the Łukasiewicz logic to include '$True$' (or more generally, '$\xi$') that will yield a naive theory once quantifiers are included.[22] Here's one well known example (from Restall 1992). Define a

[17] This is a consequence of the Brouwer Fixed Point Theorem on spaces of form $[0,1]^X$: see Field 2008, pp 97-9.

[18] A more general (and more tentative) conjecture is that if a sentence can be handled in Łukasiewicz semantics by assigning it more than one value, then the constant function corresponding to any one of those values could emerge from the revision procedure either by varying the starting valuation for conditionals or choosing non-minimal Kripkean fixed points.

[19] $\bot$ is an absurd proposition. Recall that $True(x)$ is short for "$Proposition(x) \wedge \forall y(y \,\xi\, x)$".

[20] Letting $s_\mu$ be $h_\mu(K_2) - \frac{2}{3}$ (I drop the world parameter since it doesn't matter in the example), it's easy to see that for all $\mu$, $s_{\mu+3} = -s_\mu/8$; so each sequence $s_{\lambda+n}$ rapidly converges to 0, so the value at all limits is $\frac{2}{3}$. And once it reaches the first limit, at $\omega$, it never moves from there. This analysis is independent of the somewhat arbitrary assumptions made about the values that $h_0$ gives to conditionals.

[21] These average to $\frac{2}{3}$, so in this case the value in the new system is the average (defined as a limit) of the values in the old. While this is so for many sentences, it is not so for all. Consider a sentence $W$ equivalent to $True(\lambda W) \rightarrow \neg True(\lambda W)$. This also gets value $\frac{2}{3}$ in Łukasiewicz semantics and the corresponding constant function in the present semantics; but in the revision construction using only $\{0, \frac{1}{2}, 1\}$, the value is a function that alternates between 0 and 1 at successors and has value $\frac{1}{2}$ at limits. $W$ is also an illustration of the point made in note 13, that the averaging rule for limits, without slow corrections, would yield different (and presumably less satisfying) results than we get with slow corrections.

[22] More accurately, naive truth theory in Łukasiewicz logic is $\omega$-inconsistent, in the semantic sense: arithmetically standard models of the ground language can't be consistently expanded to include naive truth in the logic. (See Restall 1992.) If we build into the naivety condition that the

function $F(n,x)$ from natural numbers and properties to properties, with $F(0,x)$ being $\lambda x(\bot)$ and for each $n$, $F(n+1,x)$ being $\lambda x[x\,\xi\, x \to x\,\xi\, F(n,x)]$. By an easy induction we get that for any $n$, and any $x$ in the domain and world $w$, $|x\,\xi\, F(n,x)|_w$ is $min\{1,\, n\cdot(1-|x\,\xi\, x|_w)\}$; so it's 1 iff $|x\,\xi\, x|_w \leq 1-\frac{1}{n}$. Let $G(x)$ be $\lambda x[\exists n(x\,\xi\, F(n,x))]$; it follows from the previous (together with an arithmetic standardness assumption: see previous footnote) that for any $x$ and $w$, $|G(x)|_w$ is 1 if $|x\,\xi\, x|_w < 1$, and 0 if $|x\,\xi\, x|_w = 1$. Now let $R$ be $\lambda z G(z)$. The above requires that $|G(R)|_w$ is 1 iff $|R\,\xi\, R|_w < 1$, hence $|G(R)|_w \neq |R\,\xi\, R|_w$; but naivety requires that $|R\,\xi\, R|_w = |G(R)|_w$, so naivety can't be achieved in Łukasiewicz semantics.

How is this handled on the "slow correcting" revision-theoretic semantics? It's not hard to show that the value of $R\,\xi\, R$ at each world at an ordinal goes in $\omega^2$-length cycles. Each one starts with the $\omega$-sequence $\langle\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, ...\rangle$, followed by $\omega$ more $\omega$-sequences $\langle 0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, ...\rangle$.[23] (The discontinuity at limits is possible because $R\,\xi\, R$ isn't equivalent to a conditional, but to something like an infinite disjunction of conditionals; the "disjuncts" $R\,\xi\, F(n,R)$ are continuous at limits.)[24]

In summary: there is no general Łukasiewicz specification for truth or property-instantiation; Łukasiewicz semantics only tells us the value that *certain* sentences would have to have in order that naivety "hold locally" for them.[25] If a sentence is such that we can get naivety to hold locally in Łukasiewicz semantics, then the modified revision semantics will assign it a constant function. So the modified revision semantics is, in a sense, "Łukasiewicz done better": it is a coherent proposal for how to deal with truth and property instantiation, that yields essentially what Łukasiewicz semantics yields where that works, but expands the value space to handle cases where Łukasiewicz fails.[26] And to repeat, in a context where these "extraordinary" paradoxical sentences (all of which involve quantification essentially) are excluded, then in that context Permutation and all the other axioms and rules of Łukasiewicz logic can be used.[27]

---

induction rule extends to formulas that contain 'True' and that suitable composition principles hold, then naivety is flat out inconsistent in Łukasiewicz logic: see Hajek, Paris and Sheperson 2000.

[23]If the starting valuation assigns 1 to every conditional then the initial $\omega^2$-cycle starts with an $\omega$-sequence of 1's, but the rest of that $\omega^2$-cycle and all later ones are the same as if all conditionals were given starting value $\frac{1}{2}$.

[24]If $\lambda$ is a limit ordinal and $m$ a fixed finite number, then as $n$ increases, $|R\,\xi\, F(n,R)|_{\lambda+m}$ increases from 0 (at $n=0$) to a maximum value; that value is $1-\frac{1}{2^m}$ (reached at $n=m$) whenever $\lambda$ isn't a multiple of $\omega^2$, and $1-\frac{1}{2^{m+1}}$ (reached at $n=m+1$) when it is a multiple. So this is also the value of $|R\,\xi\, R|_{\lambda+m}$. And since this approaches 1 as $m$ increases, it's clear that for fixed $n$, $|R\,\xi\, F(n,R)|_{\lambda+m}$ approaches 0 as $m$ increaches. So by continuity for conditionals, $|R\,\xi\, F(n,R)|_{\lambda+\omega}$ is 0 for each $n$, which is why $|R\,\xi\, R|_{\lambda+\omega}$ is 0. At limits that are multiples of $\omega^2$ the limsup of the values $|R\,\xi\, F(n,R)|_\alpha$ for fixed $n>0$ as $\alpha$ approach the limit is at least $\frac{1}{2}$ (indeed at least $1-\frac{1}{2^n}$, since setting $m=n$ above, we see that that's the value at any predecessor of form $\chi+\omega+n$). And the liminf is 0, so the value at multiples of $\omega^2$ is $\frac{1}{2}$.

[25]I.e., hold for those sentences and other closely related ones. I won't bother to make this precise.

[26]Rossi forthcoming proposes a less unified way of generalizing Łukasiewicz semantics: he takes the value space to be the union of [0,1] with a set $L$ of sets of equations. Also, he takes the valuation space to be partial: even after enriching to sets of equations as values, not every sentence gets a value.

[27]The Permutation rule lacks obviousness in a naive theory, since no such theory can have Importation ($A \to (B \to C) \models A \wedge B \to C$), whereas the obvious argument for Permutation depends on Importation and its converse. But deductively, Permutation is of immense convenience,

## 4. "Ordinary" conditionals in naive theories without $\rightarrow$

**4.1. "Ordinary" conditionals without '$\xi$' or '$\rightarrow$': 3-valued case.** Let's go back to conditionals like

If I run for President I'll win,

which need to be understood in terms of a conditional $\triangleright$ other than $\supset$ or $\rightarrow$ (assuming we're not happy to regard them as true). A Łukasiewicz-like semantics would be totally inappropriate for $\triangleright$. In the first place, the Łukasiewicz conditional is contraposable, but ordinary conditionals aren't: from the likely fact that if Sarah Palin runs for President she won't win, it doesn't follow that if she wins she won't have run. More generally, the Łukasiewicz conditional reduces to $\supset$ when the antecedent and consequent take on classical values, and we don't want that for $\triangleright$.

The most developed proposal for how $\triangleright$ might work, in contexts that *don't* involve paradoxical sentences, is in terms of a variably strict conditional in the general ballpark of Stalnaker 1968. This involves adding a "neighborhood structure" to the space $W$ of worlds, perhaps given by a ternary relation $x \leq_w y$ meaning intuitively that the change from $w$ to $x$ is no bigger than the change from $w$ to $y$. For each $w$, the binary relation $x \leq_w y$ is required to be transitive, and reflexive on its field (i.e. $\forall x \forall y(x \leq_w y \vee y \leq_w x \supset x \leq_w x)$; that field is thus $\{x : x \leq_w x\}$, and it is natural to identify this with the set $W_w$ of worlds accessible from $w$ used in the semantics for $\Box$. A simple version of the semantics in the 2-valued case is:

**(VS$_{simple}$):** $|A \triangleright B|_w$ is

$$\begin{cases} 1 & \text{if either } \neg(\exists y \in W_w)(|A|_y = 1) \text{ or } (\exists y \in W_w)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 1)] \\ 0 & \text{otherwise.} \end{cases}$$

But that is the appropriate form only if you assume, with Stalnaker, that for any $w$, and any $x$ and $y$ in $W_w$, either $x \leq_w y$ or $y \leq_w x$. If we don't make that Connectivity assumption then the appropriate version in the 2-valued case (see Burgess 1981 and Lewis 1981) is

**(VS$_{general}$):** $|A \triangleright B|_w$ is

$$\begin{cases} 1 & \text{if } (\forall x \in W_w)[|A|_x = 1 \supset (\exists y \leq_w x)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 1)]] \\ 0 & \text{otherwise.} \end{cases}$$

This reduces to (VS$_{simple}$) when the Connectivity assumption is made.

In a 2-valued context, the definition of validity will be that an inference is valid if it preserves value 1 at all normal worlds. (This is my preferred definition in multi-valued contexts too, though I will be neutral about the "value 1" part as much as possible.) I will now strengthen the first of the two structural assumptions about modal models made in Section 2.3: instead of just that for all normal worlds $w$, $w \in W_w$, I assume also that for all normal worlds $w$ and all $x \in W_w$, $w \leq_w x$. ("Weak centering at normal worlds".) This guarantees that modus ponens for $\triangleright$ is valid.

Even prior to introducing $\xi$ into the language, there might be a motivation for moving to a semantics with at least three values. For if there are worlds arbitrarily close to $w$ where $|A|$ and $|B|$ are both 1, and other worlds arbitrarily close to $w$ where $|A|$ is 1 and $|B|$ is 0, it is rather natural to think that $A \triangleright B$ should have

as anyone who works through Sections C and D of Schechter 2005 will see. So I think there is significant value in extending its range.

value $\frac{1}{2}$ at $w$, rather than the value 0 delivered by either version of (VS). So it seems natural to keep the 1 clause of (VS) (except perhaps for the decision that "vacuous conditionals", where $\neg(\exists y \in W_w)(|A|_y = 1)$, are to have value 1), but tighten the 0 clause and give value $\frac{1}{2}$ to the remaining cases. In particular, I'd suggest

**(MVS$_{simple}$):** $|A \rhd B|_w$ is

$$\begin{cases} 1 & \text{if } (\exists y \in W_w)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 1)] \\ 0 & \text{if } (\exists y \in W_w)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 0)] \\ \frac{1}{2} & \text{otherwise,} \end{cases}$$

for when Connectivity is assumed, or without that assumption

**(MVS$_{general}$):** $|A \rhd B|_w$ is

$$\begin{cases} 1 & \text{if } (\exists y \in W_w)(|A|_y = 1) \text{ and} \\ & \quad (\forall x \in W_w)[|A|_x = 1 \supset (\exists y \leq_w x)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 1)]] \\ 0 & \text{if } (\exists y \in W_w)(|A|_y = 1) \text{ and} \\ & \quad (\forall x \in W_w)[|A|_x = 1 \supset (\exists y \leq_w x)[|A|_y = 1 \wedge (\forall z \leq_w y)(|A|_z = 1 \supset |B|_z = 0)]] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

(MVS$_{general}$) reduces to (MVS$_{simple}$) when Connectivity is assumed. The decision to let vacuous conditionals have value $\frac{1}{2}$ is a matter of convenience that we should feel free to reconsider later.

It's going to be important to introduce $\rhd$ into a language with '$\rightarrow$' as well as '$\xi$', but I'll save that for Section 4.4. Without the '$\rightarrow$', adding '$\xi$' (or 'True') requires adding the value $\frac{1}{2}$ if we didn't have it already, but doesn't create a strong reason for using [0,1]:[28] that's because the variably strict semantics (MVS) used for $\rhd$ is different in character from the Łukasiewicz semantics. Nonetheless, since we will ultimately want values in [0,1], it's useful to see how they might work independent of '$\rightarrow$' and even of '$\xi$'. I'll treat that next, and get back to the paradoxes only in Section 4.3.

### 4.2. "Ordinary" conditionals without '$\xi$' or '$\rightarrow$': [0,1]-valued case.

Imagine that (perhaps for treating vagueness) we start from a model where arbitrary values in [0,1] can be assigned to atomic sentences at worlds. How (independent of any issues about 'True' and '$\xi$' and '$\rightarrow$') are we to evaluate conditionals whose antecedents and consequents are [0,1]-valued? There are several slightly different ways to do it, but they have a common theme. I'll select one that seems natural to me. For simplicity of formulation, I'll introduce the notions of lub* and glb*, which for nonempty sets are least upper bound and greatest lower bound, but where the lub* and glb* of the empty set are both $\frac{1}{2}$.

For simplicity I'll concentrate on the case where the modal model obeys Connectivity (though I will cover the case where Connectivity is not assumed in passing, starting with the next footnote). For any world $w$ and any $y \in W_w$ and any (parameterized) sentences $A$ and $B$, let

$$Lowerlim_{w,A,y}(B) =_{df} glb^*\{|B|_z : z \leq_w y \wedge |A|_z = 1\}.$$

[28] There might be some reason to invoke [0,1] even without '$\xi$': one might want to assign to each world $w$ a measure on $W_w$, and evaluate $A \rhd B$ using it together with the ordering $\leq_w$, with the idea being that the proportion of nearby $A$-worlds that have a give value for $B$ is important. But that would further complicate the discussion, so I will not pursue it.

(In an obvious terminology, $Lowerlim_{w,A,y}(B)$ is the glb* of $|B|_z$ for "$A$-worlds $z$ in the $w$-neighborhood generated from $y$".) Note that if $y_1$ is an $A$-world and $y_1 \leq_w y_2$ then for any $B$, $Lowerlim_{w,A,y_1}(B) \geq Lowerlim_{w,A,y_2}(B)$. For any $w$ and $A$ and $B$, let

$$Liminf_{w,A}(B) =_{df} lub^*\{Lowerlim_{w,A,y}(B) : y \in W_w \wedge |A|_y = 1\}$$

Roughly, it's the largest number $r$ such that there's a $w$-neighborhood that contains $A$-worlds and where at all $A$-worlds in it, $|B|$ is at least $r$.[29] Then a *partial* generalization of both (VS$_{simple}$) and (MVS$_{simple}$) is

**(CV-special):** $|A \triangleright B|_w = 1$ iff $Liminf_{w,A}(B) = 1$.[30]

This is only a partial account: it only tells us when a conditional gets value 1. For a full account, we have to decide whether it is (VS) or (MVS) that we want to generalize to the $[0, 1]$ case.

If it's (VS) that we want to generalize, we set $|A \triangleright B|_w = Liminf_{w,A}(B)$ (or if Connectivity is not assumed, $LOWERVAL_{w,A}(B)$: see note 30). For later reference, I'll use the label (CV-B) for this proposal, or rather, for this proposal modified to give value 1 to vacuous conditionals. However, I'm inclined to think it more desirable to generalize (MVS). In that case, we introduce notions dual to the previous ones:

$$Upperlim_{w,A,y}(B) =_{df} lub^*\{|B|_z : z \leq_w y \wedge |A|_z = 1\}, \text{ and}$$
$$Limsup_{w,A}(B) =_{df} glb^*\{Upperlim_{w,A,y}(B) : y \in W_w \wedge |A|_y = 1\}.$$

Then my proposed generalization of (MVS) when Connectivity is assumed is

**(CV):** $|A \triangleright B|_w$ is

$$\begin{cases} Liminf_{w,A}(B) & \text{when that is at least } \frac{1}{2}; \\ Limsup_{w,A}(B) & \text{when that is at most } \frac{1}{2}; \\ \frac{1}{2} & \text{when } Liminf_{w,A}(B) < \frac{1}{2} < Limsup_{w,A}(B). \end{cases}$$

A feature that I find attractive is that the value of $\neg(A \triangleright B)$ is the same as that of $A \triangleright \neg B$. (If we used value 1 instead of $\frac{1}{2}$ for the vacuous case, we'd need an exception to this feature for vacuous $A$.) I don't insist on the details of rule (CV), but it will serve as a good illustration for an account of ground-level conditionals that a naive theory should extend.

Everything is the same when Connectivity is dropped except that we use the $LOWERVAL_{w,A}$ of note 30, and its obvious dual $UPPERVAL_{w,A}$, instead of $Liminf_{w,A}$ and $Limsup_{w,A}$.

This, to repeat, is a natural generalization of variably strict semantics for '$\triangleright$', prior to adding '$\xi$' or '$\rightarrow$'. I now turn to what happens when we add '$\xi$', saving until Section 4.4 what happens when we add '$\xi$' and '$\rightarrow$' together.

[29] More accurately: it's the largest number $r$ such that for any $\epsilon > 0$, there's a $w$-neighborhood that contains $A$-worlds and where at all $A$-worlds in it, $|B|$ is at least $r - \epsilon$.

[30] If we don't assume that the $\leq_w$ relation obeys Connectivity, then in the [0,1] framework, we first add an argument $x$ to the notion of $Liminf$, and define a notion $LOWERVAL$ from that:

$$Liminf_{w,A,x}(B) =_{df} lub^*\{Lowerlim_{w,A,y}(B) : y \leq_w x \wedge |A|_y = 1\}.$$
$$LOWERVAL_{w,A}(B) =_{df} glb^*\{Liminf_{w,A,x}(B) : x \in W_w \wedge |A|_x = 1\}.$$

(When Connectivity is assumed, $LOWERVAL_{w,A}$ just is $Liminf_{w,A}$.) The analog of (CV-special) is then

**(CV$_{gen}$-special):** $|A \triangleright B|_w = 1$ iff $LOWERVAL_{w,A}(B) = 1$.

4.3. **"Ordinary" conditionals with '$\xi$' but not '$\rightarrow$': two approaches.** How should we handle '$\triangleright$' in the presence of '$\xi$' when we don't have to worry about '$\rightarrow$'? I'll discuss two approaches: a revision-theoretic approach, and an approach that generalizes the fixed point construction of Brady 1983. (As mentioned in note 11, there's also a "higher order fixed point" approach, but it is more complicated than the revision and doesn't appear to have any compelling advantages.) The revision-theoretic approach doesn't quite keep to the letter of (CV), but gives something in the ballpark, and reduces to (CV) for propositions without '$\xi$'. For Brady it's similar (though the Brady approach is more natural in a context where (CV-B) rather than (CV) is the target). Unfortunately, the most straightforward version of the Brady-like approach leads to some rather undesirable results: for instance, if $\top$ is a vacuous conditional-free proposition such as $\lambda\forall x(x = x)$, and $\perp$ is its negation, then $\neg(\top \triangleright \perp)$ comes out valid but$\top \triangleright \neg(\top \triangleright \perp)$ doesn't; indeed the negation of the latter comes out valid. But I'll discuss a modified version that seems to avoid the undesirable results. I will not decide among the approaches, and there could well be alternatives preferable to both.

4.3.1. *The revision approach.* On the revision approach we can be brief, because the situation is much like the revision approach for $\rightarrow$, simplified in that there is no need for slow corrections. The main difference is that here the revision procedure doesn't operate world by world, but instead operates on the assignment of values to all worlds at once.

Suppose we're given a [0,1]-valued modal model (including neighborhood structure) for the language with '$\triangleright$' but not '$\rightarrow$' or '$\xi$', and with no property or proposition abstracts. Now add property and proposition abstracts, and '$\xi$'. Then if $j$ is a "hypothesis" assigning values in [0,1] to all parameterized $\triangleright$-conditionals in the language at each world, the Kripkean procedure of Section 2 will generate a fixed-point value $|A|_{j,w}$ based on $j$, for every parameterized sentence $A$ of the language and every world $w$. And then we can use this to come up with a revised valuation $S(j)$ for $\triangleright$-conditionals, in analogy with (CV) (or the generalized version that avoids the Connectivity assumption). That is, in the version that assumes Connectivity, it will be

**(REV):** $S(j)(A \triangleright B, w)$ is

$$\begin{cases} Liminf_{w,A,j}(B) & \text{when that is at least } \frac{1}{2}; \\ Limsup_{w,A,j}(B) & \text{when that is at most } \frac{1}{2}; \\ \frac{1}{2} & \text{when } Liminf_{w,A,j}(B) < \frac{1}{2} < Limsup_{w,A,j}(B). \end{cases}$$

(The extra subscript $j$ on Liminf and Limsup is for the hypothesis that assigns values to $A$ and $B$ at each world.) Then starting with a transparent initial valuation of conditionals at worlds, we use this rule to give valuations at successor ordinals. At limit ordinals we proceed as with $\rightarrow$: use the liminf of the values at prior ordinals when that's at least $\frac{1}{2}$, the limsup when it's at most $\frac{1}{2}$, and $\frac{1}{2}$ in other cases.

Here too, the general theory of revision sequences applies, to yield a non-empty set $FIN^*$ of ordinals after which only recurring hypotheses occur.[31] And as before,

[31]The * is simply to emphasize that the set of final ordinals in the $\triangleright$-construction needn't be the same as in the $\rightarrow$-construction. For the same reason I'll use different Greek letters for the reflection ordinals of the $\triangleright$-construction.

we single out the "reflection ordinals" for this construction as distinguished members. Every recurring hypothesis appears between any two reflection ordinals, and the value of a conditional at a reflection ordinal $\Omega$ is

**(#):** $j_\Omega(A \triangleright B, w)$ is the greatest lower bound of $\{j(A \triangleright B, w) : j \text{ is recurrent}\}$, when that is at least $\frac{1}{2}$; the least upper bound when that is no more than $\frac{1}{2}$; and $\frac{1}{2}$ in all other cases.

Since valuations at successors of reflection ordinals are recurrent, this together with the revision rule gives us "one direction of" (CV): if $|A \triangleright B|_{w,\Omega} > \frac{1}{2}$ then $Liminf_{w,A,\Omega}(B) \geq |A \triangleright B|_{w,\Omega}$, and if $|A \triangleright B|_{w,\Omega} < \frac{1}{2}$ then $Liminf_{w,A,\Omega}(B) \leq |A \triangleright B|_{w,\Omega}$. (So for instance if $|A \triangleright B|_{w,\Omega} = r > \frac{1}{2}$ then for any $\epsilon > 0$, there is a $y_\epsilon$ in $W_w$ such that $(\forall z \leq_w y_\epsilon)(|A|_{z,\Omega} = 1 \supset |B|_{z,\Omega} \geq r - \epsilon)]$ , and dually when $|A \triangleright B|_{w,\Omega} = r < \frac{1}{2}$.) The reverse inequalities of (CV) do not in general hold at reflection ordinals (that is, we can have $Liminf_{w,A,\Omega}(B) > |A \triangleright B|_{w,\Omega} > \frac{1}{2}$) and $Liminf_{w,A,\Omega}(B) < |A \triangleright B|_{w,\Omega} < \frac{1}{2}$), because of the quantification over all recurrent valuations; but they do of course hold for conditionals that get the same value at every ordinal in $FIN^*$, which includes all conditionals with no occurrences of '$\xi$'.

When there is no '$\rightarrow$' in the language, we can generalize (#) from conditionals to arbitrary sentences: $|A|_{\Omega,w}$ is the greatest lower bound of the $|A|_{\alpha,w}$ for $\alpha$ in $FIN^*$, when that's at least $\frac{1}{2}$; the least upper bound when that is no more than $\frac{1}{2}$; and $\frac{1}{2}$ in all other cases. This is an analog of the "Fundamental Theorem" above, but this time for $\triangleright$-sentences. However, the result does not hold in full generality with '$\rightarrow$' in the language: it still holds for $\triangleright$-conditionals, but the inductive argument extending it to other sentences in the language is blocked, and there are sentences where $\triangleright$ is inside the scope of an $\rightarrow$ where the result fails. I don't think this is a serious problem for the revision approach: whereas the Fundamental Theorem for $\rightarrow$-sentences is important in establishing some important laws, as we'll see in Section 4.5, there is no obvious such need for a corresponding result for $\triangleright$-sentences. Nonetheless, the lack of such a theorem makes the theory less tidy, and might be seen as an advantage of the "revised general Brady construction" that I consider next.

Note that in the revision construction we've left the modal model structure completely unchanged. What is changed is the valuation space: it is now of form $[0, 1]^\Psi$, where $\Psi$ is the set of values from one reflection ordinal to the next (including at least one of them). We take an inference from $\Gamma$ to $B$ to be valid if for every modal model $M$, and every normal world in $W_M$ (which may be every world in $W_M$, depending on the modal model), if all members of $\Gamma$ have the constant function $\mathbf{1}$ as values, so does the conclusion.

4.3.2. *The Brady-like and revised Brady-like fixed point approaches.* Ross Brady introduced a procedure for naive conditionals, for 3-valued models in which $W$ consisted of only a single world so that the "variably strict" element is lost. But it easily generalizes to multi-world 3-valued models $W$, and indeed to multi-world continuum-valued models. I'll now state the idea in a way that incorporates both generalizations; several paragraphs from now I'll add a substantial tweak. In the Brady construction with or without the tweak, it's best to think of vacuous conditionals as having value 1, and probably to build upon the asymmetric rule (CV-B) of Section 4.2 rather than the symmetric (CV). But for present purposes we need not fuss about these details.

The simplest way to view the generalized Brady construction is to leave the original modal model structure unaltered, but to alter the revision rule $S$; instead, let $S^B(j)(A \triangleright B, w) = min\{S(j)(A \triangleright B, w), j(A \triangleright B, w)\}$ where $S$ is much as in the revision approach (though probably modeled on (CV-B) instead of (CV)). What I want to focus on is the minimization used in $S^B$. This modification guarantees that $S^B(j)(A \triangleright B, w) \leq j(A \triangleright B, w)$, for each $A \triangleright B$ and $w$; so that this "revision rule" is monotonic. At limit ordinals, we set $j_\lambda(A \triangleright B, w) = glb\{j_\alpha(A \triangleright B, w) : \alpha < \lambda\}$. As a result, the entire construction reaches a fixed point. We take validity to be preservation of value 1 at the fixed point, at all normal worlds in all models.

I haven't specified the question of the starting valuation $j_0$. Brady's approach was to take this as assigning the value 1 to every conditional in the one world in his model. The obvious generalization of to the multi-world context is to let it assign value 1 to every conditional at every world. But with or without this generalization to the multi-world context, this has very odd results (and would even in the original 3-valued setting). For instance, at the single world of Brady's actual approach, while $\top$ and $\neg(\top \triangleright \bot)$ quite properly get fixed point values 1, $\top \triangleright \neg(\top \triangleright \bot)$ gets fixed point value 0. That's because at the initial stage, $\neg(\top \triangleright \bot)$ gets the wrong value 0, and though this is corrected at the next stage, its effects survive: it makes $\top \triangleright \neg(\top \triangleright \bot)$ get value 0 at that next stage, and once a conditional gets value 0 it can never recover. The oddity carries over to the multi-world generalization: indeed, now $\top \triangleright \neg(\top \triangleright \bot)$ gets fixed point value 0 *at every world.*

This feature of the Brady approach is due to the starting valuation. Can we find a better starting valuation that doesn't have this feature? For a long time I didn't think one could do so without introducing ideas foreign to his approach, but it now occurs to me that there is a way; it involves adapting the suggestion for a starting valuation for '$\rightarrow$' that I made in note 16. Let $j$ be an arbitrary transparent valuation of (parameterized) conditionals at worlds; then a Brady construction that takes $j$ as its starting valuation yields as its fixed point a new valuation $Reg(j)$ of conditionals at worlds (with $[Reg(j)](A \triangleright B, w) \leq j(A \triangleright B, w)$). Now, for any set $V$ of transparent valuations, let $j^{\#(V)}(A \triangleright B, w)$ be the least upper bound of the $[Reg(j)](A \triangleright B, w)$ for all $j$ in $V$. Then a simple version of the tweak is to take as starting valuation $j^{\#(V_0)}$, where $V_0$ is the set of transparent valuations. That's enough to avoid at least the most obvious problems of the construction without the tweak. (A natural further improvement is to use instead $j^{\#(V_{fp})}$, where $V_{fp}$ is constructed by a natural fixed point procedure given in the attached footnote.)[32] Call the use of a starting valuation based on a $j^{\#}$ (whether applied to $V_0$ or to $V_{fp}$) a *revised general Brady construction.* ("General" reflecting both the use of a multi-world starting point with variably strict semantics and the use of [0,1]. The *unrevised* general Brady construction is the one starting from the valuation that assigns value 1 to every conditional at every world.)

$A \triangleright B$ won't be value functional in the values in [0,1] that $Reg(j^{\#})$ assigns at worlds: to compute its value you need the values of $A$ and $B$ at worlds supplied earlier in the fixed point construction from $j^{\#}$. But the construction does make

[32] For each $\alpha$ let $V_{\alpha+1}$ be the set of valuations $j$ where for each $A \triangleright B$ and each $w$, $j(A \triangleright B, w) \leq j^{\#(V_\alpha)}(A \triangleright B, w)$; repeated application (taking intersections at limits) gives rise to a decreasing sequence of $V_\alpha$, non-empty at each stage (as can be seen from the fact that it includes the valuation that assigns every conditional value 0 at all worlds, though of course it also includes much better ones).

it value functional in the function space $[0,1]^X$, where $X$ is the set of the fixed point ordinal and its predecessors. So this kind of product space turns up on the Brady-based construction as well as on the revision. As with the unrevised Brady, validity is taken to be the preservation of value 1 at the fixed point ordinal at all normal worlds.

I haven't investigated this construction closely, but it looks to me as if it and the revision construction do about equally well in the respects I most care about. They will not deliver exactly the same laws, but as we'll see in Section 4.5, each when combined in the right way with $\rightarrow$ leads to the most obviously desirable laws about how restricted quantification and $\triangleright$ interact; and I'm not in a position now to systematically evaluate the laws on which they differ. On simplicity of use there are tradeoffs: often the values in revised Brady are heavily dependent on the starting valuation $j^{\#}$, and calculating that adds a layer of complexity; on the other hand, in revised Brady certain conditionals can be easily seen to have value 0 at the fixed point whatever their starting values. (An informed decision on simplicity/convenience as well as on laws could only be made after we've added $\rightarrow$ to the language.)

If the revised Brady might be good as an alternative to the revision account for $\triangleright$, might it also be good as an alternative to the revision account for '$\rightarrow$'? No. For it's clear that an adequate account of a restricted quantifier conditional '$\rightarrow$' requires the Weakening Rule

**$\rightarrow$-WEAKENING:** $B \models A \rightarrow B$;

otherwise, the obviously-desirable inference from "Everything is $B$" to "All $A$ are $B$" wouldn't come out valid. And this law fails on even the revised Brady for $\rightarrow$, though less obviously so than for the unrevised. The inference from $\neg W$ to $W \rightarrow \neg W$, where $W$ is equivalent to $True(\langle W \rangle) \rightarrow True(\langle \neg W \rangle)$, is no longer a counterexample: $j^{\#}$ (in either version) in fact assigns it the same value $\frac{2}{3}$ to $W$ that it gets with Łukasiewicz, and it retains this value all the way through to the fixed point. But a counterexample can be obtained from the Restall sentence $R$ considered earlier, which (in very sloppy notation) is equivalent to $\exists n(R \rightarrow^n \bot)$. It isn't hard to see that $j^{\#}$ (in either version) assigns to $R \rightarrow^n \bot$ the value $\frac{n}{n+1}$, and that at the $\alpha^{th}$ stage its value is 0 when $\alpha \geq n$ and $\frac{n-\alpha}{n-\alpha+1}$ otherwise. The value of $R$ at any stage is the least upper bound of the values of these conditionals, so it's 1 at all finite stages, and 0 from stage $\omega$ on. From this it's clear that $\top \rightarrow \neg R$ gets value 0 at least from stage 1. Consequently, while $\neg R$ gets value 1 at the fixed point, $\top \rightarrow \neg R$ gets value 0, in violation of Weakening.[33]

An analogous sentence using $\triangleright$ for $\rightarrow$ would of course show the invalidity of a Weakening rule for $\triangleright$, but that's no problem since for $\triangleright$ we don't expect Weakening to hold. (We might expect $\Box B \models A \triangleright B$, but I don't regard its failure to hold unrestrictedly on the revised Brady as obviously crippling.)

[33][Added at the last minute.] One might consider changing the notion of validity, to preserving the property of having value 1 *throughout the fixed point construction.* On the untweaked Brady construction that would validate Weakening, but would be obviously unsatisfactory: no negation of a conditional would be valid. With the new starting valuation that objection doesn't apply, but it is also unobvious that Weakening holds. ($\neg R$ is no longer a counterexample, but I've been unable either to prove there are no others or to find any.) If Weakening is valid on this definition, and if the recapture of Łukasiewicz values that we get for sentence $W$ extends to all sentences in the quantifier-free sublanguage, then this rival approach to $\rightarrow$ is worth seriously considering.

4.4. $\triangleright$ **and** $\rightarrow$ **together.** If we want to employ $\triangleright$-conditionals in a setting with naive truth or property-instantiation, we need to deal with sentences that contain $\rightarrow$ and $\triangleright$ together. Indeed, many obvious laws of conditionals have $\rightarrow$ embedded inside $\triangleright$: e.g. $[\forall x Bx \wedge \forall x(Bx \rightarrow Cx] \triangleright \forall x Cx$, i.e. "If everything is $B$, and all $B$ are $C$, then everything is $C$". (There are fewer obvious laws that essentially require embedding $\triangleright$ inside $\rightarrow$, but there are some: for instance, $True(\lambda(A \triangleright B)) \leftrightarrow [True(\lambda A) \triangleright True(\lambda B)]$.)

We have the pieces, but need to combine them, and do so in a way that will yield desirable laws. The right way to do so is asymmetric between $\triangleright$ and $\rightarrow$ (as it would pretty much have to be if we decide on the Brady option for $\triangleright$, since it's hard to see how to combine that symmetrically with a very different kind of account for $\rightarrow$).

Suppose we're given a modal model for the ground language, taken now to include the neighborhood structure given by the relative closeness ordering in worlds. We want to use it to assign values to all parameterized sentences in the language with '$\xi$', '$\rightarrow$' and '$\triangleright$' (plus abstraction terms, 'Property' and 'Proposition').

In either case, the overall procedure I propose involves a multi-stage *macro-construction*, focused on $\triangleright$; each stage of which is a *mid-level construction*, focused on $\rightarrow$; each stage of which is a (generalized) Kripkean *micro-construction*, focused on '$\xi$'.

In the microconstructions, we hold fixed both a [0,1]-valuation $j$ of parameterized $\triangleright$-conditionals at each world and a [0,1]-valuation $h$ of parameterized $\rightarrow$-conditionals at each world, and use the generalized Kripke fixed point construction in Section 2 to get a value in [0,1] for every parameterized sentence of the language at every world relative to $j$ and $h$.

In the mid-level constructions, we hold fixed a [0,1]-valuation $j$ of parameterized $\triangleright$-conditionals at each world, and use the revision procedure of Section 3 to get a value $||A||_{j,w}$ for every sentence at each world. These values $||A||_{j,w}$ are in a space $[0,1]^{\Pi_j}$ where $\Pi_j$ is the difference between two reflection ordinals of the $\rightarrow$-construction based on $j$.[34] As we've seen, for typical sentences $A$ including finitely iterated Curry sentences, the value $||A||_{j,w}$ for any $j$ and $w$ will be a constant function on $\Pi_j$: that was the advantage of using [0,1] instead of $\{0, \frac{1}{2}, 1\}$. But for weird enough sentences, it won't be. In that case, though, the value at the reflection ordinal has a privileged status, which will be exploited in the macro-construction. (The values that sentences take on at other final ordinals of the midlevel construction won't be directly used in the macro-construction.)

In the macro-construction, we vary the [0,1]-valued valuation $j$ of parameterized $\triangleright$-conditionals at worlds, either by the revision procedure described in 4.3.1 or the revised Brady procedure described in 4.3.2. The only difference between here and the previous section is that the $\triangleright$-conditionals that are evaluated in the stages of the macro-construction may now contain $\rightarrow$, but this is no problem since the mid-level reflection values of all sentences including ones with $\rightarrow$ get values at prior stages of the macro-construction.

So the overall architecture of the two constructions is the same. In both cases we can think of the overall valuation space as rather like a fiber bundle. The "base space" $Z$ is a segment of the ordinals, with a distinguished member $z_0$ (the reflection ordinal or fixed point ordinal of the macro-construction, as the case may be).

[34] If we like, we can get a common space $[0,1]^{\Pi}$ that works for all $\triangleright$-valuations $j$. ($\Pi$ need only be a common right-multiple of each of the $\Pi_j$. A big enough initial ordinal will certainly do.)

To each member of $Z$ is attached a circular "fiber", obtained from $[\Delta_j, \Delta_j + \Pi_j]$ by identifying endpoints, which is attached to the base space at its distinguished member $\Delta_j$. In the evaluation of sentences, values in [0,1] are assigned to $\triangleright$-conditionals at worlds primarily at the base points (though the value is used at all points of the fiber attached to the base point); whereas values of $\rightarrow$-conditionals at worlds are assigned primarily to points of the various fibers. Moreover, the rules for evaluating sentences on the fibers are the same in both constructions: they're given by the mid-level procedure. The only difference is on the valuation rules in the base space.

It might well be thought that the "fiber bundle" structure that these constructions have in common is implausibly complicated. But I know of no other way of adequately accommodating the two kinds of conditionals in a naive theory. (I'll briefly consider another approach in Section 5.)

4.5. **Some laws.** All these constructions validate some important laws, including the following:

**(1):** $[\forall x(Ax \rightarrow Bx) \wedge Ay] \triangleright By$ "If all $A$ are $B$, and $y$ is $A$, then $y$ is $B$"
**(1c):** $[\forall x(Cx \rightarrow Dx) \wedge \neg Dy] \triangleright \neg Cy$
"If all $C$ are $D$, and $y$ is not $D$, then $y$ is not $C$"
**(2):** $\forall x Bx \triangleright \forall x(Ax \rightarrow Bx)$ "If everything is $B$, then all $A$ are $B$"
**(2c):** $\forall x \neg Cx \triangleright \forall x(Cx \rightarrow Dx)$ "If nothing is $C$, then all $C$ are $D$".
**(3):** $\forall x(Ax \rightarrow Bx) \wedge \forall x(Bx \rightarrow Cx) \triangleright \forall x(Ax \rightarrow Cx)$
"If all $A$ are $B$ and all $B$ are $C$ then all $A$ are $C$"
**(4):** $\forall x(Ax \rightarrow Bx) \wedge \forall x(Ax \rightarrow Cx) \triangleright \forall x(Ax \rightarrow Bx \wedge Cx)$
"If all $A$ are $B$ and all $A$ are $C$ then all $A$ are both $B$ and $C$"
**(5):** $\neg\forall x(Ax \rightarrow Bx) \triangleright \exists x(Ax \wedge \neg Bx)$
"If not all $A$ are $B$, then something is both $A$ and not $B$"
**(5*):** $\neg\exists x(Ax \wedge \neg Bx) \triangleright \forall x(Ax \rightarrow Bx)$
"If nothing is both $A$ and not $B$, then all $A$ are $B$"
**(6):** $\exists x(Ax \wedge \neg Bx) \triangleright \neg\forall x(Ax \rightarrow Bx)$
"If something is both $A$ and not $B$, then not all $A$ are $B$".

(1c) and (2c) follow from (1) and (2) respectively on the supposition that the $\rightarrow$ contraposes, as the $\rightarrow$ in this paper does; but they are worth stating separately for those who would like a non-contraposable $\rightarrow$. Also (5*) needs to be stated separately from (5), since the ordinary conditional $\triangleright$ definitely does not contrapose. (5*) is just a more general form of (2).

These are very simple laws: all of them are schemas of form $X \triangleright Y$, where neither $X$ nor $Y$ involve $\triangleright$ essentially though they do involve $\rightarrow$. That makes them very easy to verify.

For in order that $X \triangleright Y$ be valid, it is required only that for every modal model for the ground language and every member $j$ of the base space $Z$ (that is, the set of recurrent macro-valuations, on the revision macro-construction; and the set of stages along the way to the fixed point, on the Brady) and every normal world:

> if $|X|_{j,w,\Delta_j}$ is 1 then so is $|Y|_{j,w,\Delta_j}$ (where $\Delta_j$ is a reflection ordinal of the fiber attached to $j$).

Using the Fundamental Theorem for fibers, this is in turn equivalent to:

**(*):** for all normal $w$ and all $j \in Z$, if $|X|_{j,w,\alpha} = 1$ for all final $\alpha$ in the $\rightarrow$-construction for $j$,$w$, then $|Y|_{j,w,\Delta_j} = 1$.

In the case of each of the laws $X \triangleright Y$ above, (*) would hold even without the restriction to normal $w$, or the restriction to $j \in Z$. In other words, the validity is guaranteed by only the fiber construction together with basic structural features of the macro-construction that are common to its different versions.

For instance, in the case of law 1, what we need (dropping the $j$ and $w$ subscripts from the notation, since they are irrelevant to the argument) is that if $(\forall\alpha \in FIN)[|\forall x(Ax \to Bx) \wedge Ao|_\alpha = 1]$ then $|Bo|_\Delta = 1$, for any object $o$ in the domain. But the assumption requires that $|Ao|_\Delta$ be 1 and also that for any final $\alpha$ in the $\to$-construction, $|Ao|_\alpha \leq |Bo|_\alpha$; and since $\Delta$ is itself one of those final $\alpha$, $|Bo|_\Delta = 1$ as desired.

And in the case of 5*, what we need is that if $(\forall\alpha \in FIN)[|\neg\exists x(Ax \wedge \neg Bx)|_\alpha = 1]$ then $|\forall x(Ax \to Bx)|_\Delta = 1$. But the assumption requires that for any $o$ in the domain, $|\neg Ao \vee Bo|_\Delta = 1$, which requires that either $|Ao|_\Delta = 0$ or $|Bo|_\Delta = 1$, which by the Fundamental Theorem requires that either $(\forall\alpha \in FIN)(|Ao|_\alpha = 0)$ or $(\forall\alpha \in FIN)(|Bo|_\alpha = 1)$. And that guarantees $|\forall x(Ax \to Bx)|_\Delta = 1$, as desired.

The other laws I've listed are verified similarly.

## 5. Standard "Relevant" Conditionals.

Much of the technical literature on the paradoxes, especially that in the dialetheic tradition, is focused on relevant conditionals, especially those in the vicinity of the System **B** of Priest 2008. (In some of the literature this system is weakened to exclude even the rule form of contraposition. In some cases (with or without the modification about contraposition) it is strengthened to include excluded middle. There may be other variations as well.)

It is not entirely clear to me the motivation behind this focus. Such relevant conditionals are ill-suited both for restricted quantification and for the ordinary English conditional that we find in such sentences as

**(1):** If I get a reservation at that restaurant, I'll eat dinner there tonight.

One of the most obvious features of such sentences is that they do not obey the rule of antecedent strengthening

**(AS):** If $A$ then $C \models$ If $A$ and $B$ then $C$;

for (1) clearly doesn't imply

**(2):** If I get a reservation at that restaurant and die immediately after doing so, I'll eat dinner there tonight.

But (AS) is built into System **B** and the logics in its vicinity that the technical literature above employs.

It is no defense to say that if a logic is compatible (in the sense of Post-consistency) with naivety in a logic with (AS), then it is compatible with naivety in the weaker logic with (AS) dropped. For what we want is more than the Post-consistency of the naivety claims alone, we want that *adding naivety to whatever acceptable assumptions we start with* is Post-consistent. And since assumptions violating (AS) are acceptable, we need to demonstrate that naivety is compatible with such violations. Showing that naivety holds *in some models of a logic that includes (AS)* can't be of any help.

It is equally clear that relevant conditionals are of no help for restricted quantification. Indeed, it seems pretty clear that for that, we need a logic whose $\to$ reduces to $\supset$ when the antecedent and consequent are classical; and the relevant conditionals were designed precisely *not* to have that feature. That aside, even many who think relevant conditionals important have conceded that they can't be used for restricted quantification because we need at least the rule form of Law 2 above, viz.

Everything is $B \models$ All $A$ are $B$.

If $\rightarrow$ is the restricted quantifier conditional, the conclusion is $\forall x(Ax \rightarrow Bx)$, and so the law clearly requires $D \models C \rightarrow D$; and that law is not valid for relevant conditionals. For this reason, the authors of Beall *et al* 2006 have advocated using both a relevant conditional for purposes other than restricted quantification and an "irrelevant" one for restricted quantification. I disagree about the details of their "irrelevant" one (theirs doesn't reduce to $\supset$ in classical contexts), but they are certainly right that no relevant conditional will do for restricted quantification.

The Beall *et al* paper was a great advance: indeed my own focus on the use of two distinct conditional operators, one for restricted quantification and one for more ordinary conditionals, was inspired by it. But for the reasons just given, I don't think that either of the two particular conditional operators they use is what is needed for their respective purposes.

Further evidence for this, were it needed, is that the laws that one gets with their conditionals are very far from what we need for restricted quantification. I concede that there might be some dispute as to exactly which laws we need. Dialetheists *must* dispute my list: for instance, even *the rule forms of* (1) and (5*) together entail

Everything is either $B$ or is not $A$, $c$ is $A \models c$ is $B$;

and taking $B$ to be $x = x \wedge \bot$ and $A$ to be $x = x \wedge \lambda$ where $\lambda$ is a dialetheia, this yields

$$\neg\lambda, \lambda \models \bot,$$

which no dialetheist can accept. Beall *et al* do accept the rule form of (1), and so reject even the rule form of (5*). It's hard for me to find an independent rationale for rejecting it, short of attributing to the restricted quantifier a modal element which I don't think it has; but I recognize that I'm unlikely to convince the committed dialetheist that there is a problem here.[35]

But dialetheism aside, the Beall *et al* system also doesn't include the full (1), or (3), (5) or (6). (The failure to get the full (1) is because they assume an intimate relation between the two conditionals: they assume that $A \triangleright B \models A \rightarrow B$ (using $\triangleright$ for their relevant conditional and $\rightarrow$ for their restricted quantifier conditional), which would mean that (1) would require "Pseudo Modus Ponens" for $\rightarrow$ (i.e. $[(A \rightarrow B) \wedge A] \rightarrow B$); as is well-known, that conflicts with genuine modus ponens in any naive theory.) There's a lot more that could be said about this, but one moral seems to be that relevant conditionals are just the wrong tool for naive theories.[36]

And whatever one thinks of the treatment of $\triangleright$ in these theories, the treatment of the restricted quantifier conditional is highly problematic: though these theorists

[35] The weak consequence ($2_c$) of (5*) is also incompatible with the rule form of (1). Their restricted quantifier conditional is non-contraposable, so they can accept (2).

[36] This is perhaps a slight overstatement, in that I've suggested a possible use for the Brady construction, which is connected to relevance. But I've suggested that it has no role for the $\rightarrow$-conditional, and the main departure of $\triangleright$ from a conditional that reduces to $\supset$ in classical contexts isn't due to the Brady construction but to the variably strict semantics.

recognize that it is not a relevant conditional, they treat it as intimately related to one, in a way that prevents its collapse to $\supset$ in classical contexts.

## 6. Property identity

An important issue about naive property theory, raised originally in Restall 2010, is how to include within it satisfactory identity conditions for properties.

6.1. **A negative result.** Restall assumed that a satisfactory account of property-identity should satisfy the condition that if $P(x) \models Q(x)$ and $Q(x) \models P(x)$ then $\models \lambda x P(x) = \lambda x Q(x)$, and showed that if so, then no satisfactory account is possible in a theory like mine. But as I argued in Field 2010, that condition on property identity is unreasonable: if $\mu$ is a Liar sentence, and $P(x)$ is "$\mu \wedge x$ is a cockroach" and $Q(x)$ is "$\mu \wedge x$ is a kangaroo", we shouldn't expect $\lambda x P(x)$ to be the same as $\lambda x Q(x)$.

A more plausible criterion, for the language without $\Box$ or $\triangleright$ that was there under consideration, is this:

**(?):** If $\models \forall x(P(x) \leftrightarrow Q(x))$ then $\models \lambda x P(x) = \lambda x Q(x)$.

And Restall's impossibility proof doesn't work for this loosened criterion. Something I said in the above paper implies that (?) can be achieved for the language there under consideration, but the argument-sketch I gave for that claim was seriously flawed. Indeed, Harvey Lederman pointed out to me that on the specific version of the $\rightarrow$ there under consideration (which involved a jumpy correction rule and a starting valuation that assigned every conditional value $\frac{1}{2}$), (?) must fail for any reasonable notion of identity when $P(x)$ is $\bot$ and $Q(x)$ is $\top \rightarrow \bot$.[37]

That argument would be blocked by the choice of initial valuation recommended in note 16, or alternatively by the use of slow corrections together with a starting valuation that assigns all conditionals value 1. But any hope this might generate would be misplaced.

For Tore Fjetland Øgaard gave a proof that (?) leads to triviality in any system that includes the $\rightarrow$-Weakening rule plus minimal other laws; it's reported in Section 10 of Field, Lederman and Øgaard forthcoming (as is the earlier Lederman result). Our discussion there might suggest that Øgaard's proof undermines only the idea that *coextensiveness* suffices for identifying abstracts, but in fact it undermines even the idea that *validity of* coextensiveness is sufficient.

---

[37] If identity is to behave at all reasonably, then for any formula $S(y)$, (?) implies

**(?-$S$):** If $\models \forall x(P(x) \leftrightarrow Q(x))$ and $\models S(\lambda x Q(x))$, then $\models S(\lambda x P(x))$.

Using the $P$ and $Q$ in the text, the first antecedent holds, so this becomes

**(?$^*$-$S$):** If $\models S(q)$, then $\models S(p)$, where $p$ is $\lambda x(\bot)$ and $q$ is $\lambda x(\top \rightarrow \bot)$.

Let $S(y)$ be $\forall u(y \,\xi\, u \leftrightarrow q \,\xi\, u)$. Then $\models S(q)$. But not $\models S(p)$, at least with fast corrections and a starting valuation that values all conditionals the same. To see this, we proceed in three steps. First we argue that if $\alpha > 0$ then $|S(p)|_\alpha \leq |\top \rightarrow S(p)|_\alpha$. (That's because by the definition of $S$, $|S(p)|_\alpha \leq |p \,\xi\, \lambda z S(z) \leftrightarrow q \,\xi\, \lambda z S(z)|_\alpha$, and by naivety the right hand side equals $|S(p) \leftrightarrow S(q)|_\alpha$, and for $\alpha > 0$ $|S(q)|_\alpha = 1$.) Second, we use this to argue that on the fast-correction construction, if $S(p)$ has value less than 1 when $\alpha$ is 1 then it can't have a higher value for larger $\alpha$. Finally we observe that on the assumed starting valuation, $S(p)$ does have a value less than 1 when $\alpha$ is 1: for letting $e$ be $\lambda z[\neg\exists w(w \,\xi\, z)]$, $|p \,\xi\, e|_0 = |\neg\bot|_0 = 1$ and $|q \,\xi\, e|_0 = |\neg(\top \rightarrow \bot)|_0 < 1$ and so $|S(p)|_1 \leq |p \,\xi\, e \leftrightarrow q \,\xi\, e|_\alpha < 1$. The second of the three steps fails with slow corrections when the starting value for conditionals is 1, and the third fails for any starting valuation where $\top \rightarrow \bot$ is given value 0.

The failure of (?) might be unsurprising in the presence of non-normal worlds, since its antecedent only requires that $\forall x(P(x) \leftrightarrow Q(x))$ have value **1** at *normal* worlds, and it might well be thought that failure of coextensivity at non-normal worlds precludes property identity. That could be handled by putting a '$\Box$' or a '$\top\rhd$' before the '$\forall$', given the structural assumption that if there are non-normal worlds then each is accessible from a normal world. Will this or some similar modality $\Box^*$ (perhaps defined using '$\top \to$' as well as '$\Box$' and '$\top\rhd$') solve the problem?[38] That is, is there some modality obeying minimally reasonable laws for which we can define property identity so as to get

**($?_w$):** If $\models \Box^*\forall x(P(x) \leftrightarrow Q(x))$ then $\models \lambda x P(x) = \lambda x Q(x)$?

No: Øgaard's proof generalizes to rule that out. Though the generalization is totally routine, I include a proof in a footnote since I think our original presentation was hard to survey.[39]

6.2. **A positive result.** Despite the Øgaard proof, there is room for a great deal of coarse-graining. Precisely how coarse-grained to go is somewhat arbitrary: we can pick any formula $R(x,y)$ that satisfies the following conditions.[40]

**(I):** $\models \forall x,y[R(x,y) \to (Property(x) \wedge Property(y)) \vee (Proposition(x) \wedge Proposition(y)) \vee x =_o y]$ (where $=_o$ is the ground level identity predicate)

**(II):** $\models \forall x,y[R(x,y) \vee \neg R(y,x)]$

**(III):** $\models \forall x R(x,x)$

[38] In fact the extra operators couldn't help: while (e.g.) $\top\rhd(\top \to B)$ doesn't follow from $\Box B$, still the validity of the former follows from the validity of the latter in the logics we've discussed.

[39] Define $b \sim c$ ("congruence") to mean $\forall z(b\,\xi\,z \leftrightarrow c\,\xi\,z)$; so by transparency, we have $\models b \sim c \to (A(b) \leftrightarrow A(c))$, for any $A$. In any of these logics that in turn implies $A(b) \models b \sim c \to (\top \leftrightarrow A(c))$. (This depends on $\to$-Weakening.) Then for any reasonable modality $\Box^*$ we get the "modal quasi-substitutivity lemma" $\Box^* A(b) \models \Box^*[b \sim c \to (\top \to A(c))]$.

Then let $\emptyset$ be $\lambda x(\bot)$, and if $p$ is an instantiation-invariant property, let $F(p)$ ("the quasi-complement of $p$") be $\lambda x(p \sim \emptyset)$. Trivially, $\emptyset\,\xi\,F(\emptyset)$; so making use of the instance of quasi-substitutivity where $A(p)$ is $\emptyset\,\xi\,p$, $b$ is $F(\emptyset)$ and $c$ is $\emptyset$, we easily prove

(1) $\models \Box^*[F(\emptyset) \sim \emptyset \to \bot]$.

Consider the "Hinnion property" $H =_{df} \lambda y[\lambda x(y\,\xi\,y) \sim \emptyset]$. Let $\kappa$ be $H\,\xi\,H$, and let $B$ be $\lambda x[\kappa]$. Making use of (1) together with the instance of quasi-substitutivity where $A(p)$ is $p \sim F(p)$, $b$ is $B$ and $c$ is $\emptyset$, we easily prove

(2) $\Box^*(B \sim F(B)) \models \Box^*(B \sim \emptyset \to \bot)$.

We easily prove

(3) $\models \Box^*(\kappa \leftrightarrow (B \sim \emptyset))$;

in effect, that $B$ is necessarily coextensive with $F(B)$. And from (2) and (3) we easily prove

(4) $\Box^*(B \sim F(B)) \models \Box^*(\kappa \leftrightarrow \bot)$,

whose conclusion says in effect that $B$ is necessarily coextensive with $\emptyset$.

So far, the proof has made no use of ($?_w$). I'll assume that property identity is "rigid" in the sense that if properties $p$ and $q$ are identical then $\Box^*\forall r(p\xi r \leftrightarrow q\xi r)$ (that is, $\Box^*(p \sim q)$). So ($?_w$) entails

(?*) If $\models \Box^*\forall x(P(x) \leftrightarrow Q(x))$ then $\models \Box^*[\lambda x P(x) \sim \lambda x Q(x)]$.

Applying that first to (3), we get the premise of (4); hence we get

(5) $\models \Box^*(\kappa \leftrightarrow \bot)$.

And applying (?*) again to that, we get

(6) $\models \Box^*(B \sim \emptyset)$.

But (6) and (3) yield $\models \Box^*\kappa$, which with (5) yields $\Box^*\bot$ and hence $\bot$.

[40] It may be possible to liberalize this by dropping Condition (II), weakening (V) to

**($V_{weak}$):** $R(x,y) \models \forall z[R(y,z) \to R(x,z)]$,

and finding a suitable weakening of (VI). But the use of non-bivalent property identity raises issues that would take us too far afield.

**(IV):** $\models \forall x[R(x,y) \to R(y,x)]$
**(V):** $\models \forall x,y,z[R(x,y) \wedge R(y,z) \to R(x,z)]$
**(VI):** If $b$ and $c$ are properties or propositions such that $\models R(b,c)$, and $j$ is a $\rhd$-valuation and $v$ is an $\to$-valuation and *the pair $\langle j,v\rangle$ occurs at some Kripke fixed point in the overall construction*, then for all objects $o$ in the domain and all worlds $w$, $|o\,\xi\,b|_{j,v,w} = |o\,\xi\,c|_{j,v,w}$.

Many such $R$ are easily definable in the language (even the fragment without 'True'), as long as the ground language is adequate to syntax, or to set theory.

An obvious proposal is to take $R$ to mean provable equivalence in some suitable logic; for instance, $R[\lambda uB(u), \lambda vC(v)]$ would be of form $\vdash_{log} \forall v(B(v) \leftrightarrow C(v))$. This will automatically validate the first five conditions. The "suitable logic" is naturally taken to include quantified $S_3$ ("symmetric Kleene logic"),[41] and also to include the naivety conditions plus a selection of laws involving $\to$ and $\rhd$. The latter laws need to be chosen in a way that is compatible with condition (VI), but that allows for quite a bit. For instance, (VI) isn't violated by building in the equivalence of $\lambda u(B(u) \to C(u))$ to $\lambda u(B(u) \to C(u) \wedge C(u))$ or to $\lambda u(B(u) \to u\,\xi\,\lambda zC(z))$. It also isn't violated by building in the equivalence of $\lambda uB(u)$ to $\lambda u(B(u) \wedge (C(u) \to C(u))$, or the equivalence of $\lambda u(B(u) \to B(u))$ to $\lambda u(B(u) \rhd B(u))$, provided that we start the revision process by assigning these conditionals value 1 rather than say $\frac{1}{2}$. (So this is one place where my choice of $\frac{1}{2}$ as initial value for conditionals was sub-optimal. The choice of 1 for all conditionals, though less natural, would have been marginally better and have no obvious downside; the valuation suggested in note 16 would be better still.)

I claim that any $R$ meeting (I)-(VI) meets the formal conditions on identity: in particular, the requirement of substitutivity of identity. To show this, I use the "Micro-Extensionality Theorem" from Field *et al* forthcoming (which perhaps was implicit in some much earlier papers by Ross Brady). Or rather, I use a generalization of this theorem, not only to a modal setting but more substantially, to the setting of the Kripke algebra [0,1]. (In fact it generalizes to arbitrary Kripke-algebras, but there will be no need for that.) Let $v$ and $j$ be transparent valuation functions for $\to$-conditionals and $\rhd$-conditionals respectively. Let $b$ and $c$ be any two closed property abstracts. Call $v$ *$\langle b,c\rangle$-congruent* if for any parameterized formulas $P(u)$ and $Q(u)$ and any world $w$, $v(P(b) \to Q(b), w) = v(P(c) \to Q(c), w)$; and analogously for $j$, using $\rhd$ instead of $\to$. (And call a pair $\langle j,v\rangle$ *$\langle b,c\rangle$-congruent* if both its members are.) Call a pair $\langle j,v\rangle$ *$\langle b,c\rangle$-good* if for any object $o$ in the domain and world $w$, $|o\,\xi\,b|_{j,v,w} = |o\,\xi\,c|_{j,v,w}$, where these are the Kripke fixed point values. Call a pair $\langle j,v\rangle$ *strongly $\langle b,c\rangle$-congruent* if for any parameterized 1-formula $A(z)$ in the domain, $|A(b)|_{j,v,w} = |A(c)|_{j,v,w}$. Then we have

**Generalized Micro-Extensionality Theorem:** For any closed property abstracts $b$ and $c$, and any pair $\langle j,v\rangle$ of transparent valuations for the two kinds of conditionals: if $\langle j,v\rangle$ is $\langle b,c\rangle$-congruent and $\langle b,c\rangle$-good then it is *strongly* $\langle b,c\rangle$-congruent.

**Proof:** Suppose $\langle j,v\rangle$ is not $\langle b,c\rangle$-congruent. Then there is at least one parameterized 1-formula $A(x)$ and world $w$ such that $|A(b)|_{j,v,w} \neq |A(c)|_{j,v,w}$. Call any such pair of $A(x)$ and $w$ a *counterexample.* For any counterexample $\langle A(x),w\rangle$, either

[41] This is the 3-valued logic with Kleene evaluation rules in which for an inference to be valid, its conclusion must in every model have value at least that of the minimum of the values of its premises.

**(i):** $|A(b)|_{j,v,w} < |A(c)|_{j,v,w} \leq \frac{1}{2}$
**(ii):** $|A(c)|_{j,v,w} < |A(b)|_{j,v,w} \leq \frac{1}{2}$
**(iii):** $|A(b)|_{j,v,w} > |A(c)|_{j,v,w} \geq \frac{1}{2}$, or
**(iv):** $|A(c)|_{j,v,w} > |A(c)|_{j,v,w} \geq \frac{1}{2}$.

That is, for any counterexample $\langle A(x), w\rangle$, there are ordinals $\sigma$ such that in the Kripke microconstruction, either

**($i_\sigma$):** $|A(b)|_{j,v,w,\sigma} < |A(c)|_{j,v,w} \leq \frac{1}{2}$;
**($ii_\sigma$):** $|A(c)|_{j,v,w,\sigma} < |A(b)|_{j,v,w} \leq \frac{1}{2}$;
**($iii_\sigma$):** $|A(b)|_{j,v,w,\sigma} > |A(c)|_{j,v,w} \geq \frac{1}{2}$;
**($iv_\sigma$):** $|A(c)|_{j,v,w,\sigma} > |A(b)|_{j,v,w} \geq \frac{1}{2}$.

So for any $A(x)$ that is the first component of a counterexample, there is a smallest $\sigma$ such that for some world $w$, one of $(i_\sigma)$-$(iv_\sigma)$ holds. Call this the ***height*** of $A(x)$. (All and only those $A(x)$ that are the first components of counterexamples are assigned heights.)

Assuming there are counterexamples, there are ones whose first component has lowest height; let $\delta$ be the lowest height at which there are counterexamples.

**Lemma:** No parameterized 1-formula of form $t(x)\,\xi\, x$ can have height $\delta$.

**Proof of Lemma:** If $t(x)\,\xi\, x$ had height $\delta$, then one of the cases $(i_\delta)$-$(iv_\delta)$ would apply. Relabelling if necessary, we can stick to cases $(i_\delta)$ and $(iii_\delta)$; and the proofs for them are similar, so let's just focus on $(i_\delta)$. That is, we suppose

$$|t(b)\,\xi\, b|_{j,v,w,\delta} < |t(c)\,\xi\, c|_{j,v,w} \leq \tfrac{1}{2}.$$

$\delta$ can't be 0, since $|t(b)\,\xi\, b|_{j,v,w,0} = \frac{1}{2}$, and it can't be a limit since if $|t(b)\,\xi\, b|_{j,v,w,\sigma} \geq |t(c)\,\xi\, c|_{j,v,w}$ for all $\sigma$ less than a limit $\lambda$ then the same holds for $\lambda$. So it would have to be of form $\tau+1$. But $b$ and $c$ are of form $\lambda x B(x)$ and $\lambda w C(w)$ where $B(x)$ and $C(w)$ are parameterized 1-formulas, so $|t(b)\,\xi\, b|_{j,v,w,\delta} = |B(t(b))|_{j,v,w,\tau}$; and since by the Kripke fixed point condition $|t(c)\,\xi\, c)|_{j,v,w} = |C(t(c))|_{j,v,w}$, we have

$$|B(t(b))|_{j,v,w,\tau} < |C(t(c))|_{j,v,w} \leq \tfrac{1}{2}.$$

But since $\tau < \delta$, $B(t(x))$ is not part of a counterexample. So the left side equals $|B(t(c))|_{j,v,w\tau}$, and so

$$|B(t(c))|_{j,v,w,\tau} < |C(t(c))|_{j,v,w} \leq \tfrac{1}{2}.$$

But then monotonicity yields the following relation among the fixed point values:

$$|B(t(c))|_{j,v,w} < |C(t(c))|_{j,v,w} \leq \tfrac{1}{2}.$$

So by the fixed point condition,

$$|t(c)\,\xi\, b|_{j,v,w} < |t(c)\,\xi\, c|_{j,v,w} \leq \tfrac{1}{2}.$$

But this violates the $\langle b, c\rangle$-goodness of $\langle j, v\rangle$. ■

Continuing the proof of the theorem, we generalize the Lemma: we show that no parameterized 1-formula *of any form* can have height $\delta$. This is by induction on complexity. Any atomic parameterized 1-formula of the language, say with $x$ the free variable if there is one, is either

(A) an atomic formula of the ground language
(B) of form $Property(t(x))$ or $Proposition(t(x))$
(C) of form $t(x)\,\xi\, N$ where $N$ is a name in the ground language
(D) of form $t(x)\,\xi\, \lambda y D(x,y)$ or $t(x)\,\xi\, \lambda D(x)$
(E) of form $t(x)\,\xi\, x$.

(We allow that the parameterized terms and formulas not contain free the variables shown.) Clearly no atomic formulas of form (A)-(C) can be (components of) counterexamples of any height. By the Lemma, none of form (E) can be one of height $\delta$. So it remains only to show that none of form (D) can have height $\delta$. Again we need to divide up into four cases, but a typical one would be that for some $v$,

$$|t(b)\,\xi\,\lambda y D(b,y)|_{j,v,w,\delta} < |t(c)\,\xi\,\lambda y D(c,y)|_{j,v,w} \leq \tfrac{1}{2}.$$

But again, $\delta$ can't be $\frac{1}{2}$ or a limit, so letting $\tau$ be its immediate predecessor, we have

$$|D(b,t(b))|_{j,v,w,\tau} < |D(c,t(c))|_{j,v,w} \leq \tfrac{1}{2},$$

and so we have a counterexample of height less than $\delta$, which is a contradiction.

We've shown that no atomic parameterized formula can be (a component of) a counterexample. And it is routine to extend this to non-atomic, which we do by induction on complexity. The clauses for the ordinary connectives and quantifiers are routine, and for the $\rightarrow$ and $\rhd$ clauses we use (for the only time) the assumption that $j$ and $v$ are $\langle b,c\rangle$-congruent. ■ ■

Now let's apply the theorem. Condition (VI) says that if $b$ and $c$ are $R$-equivalent then every $\langle j,v\rangle$ pair that occurs (as a Kripkean fixed point) anywhere in the overall construction is $\langle b,c\rangle$-good. So by the theorem, any such $\langle j,v\rangle$ that is $\langle b,c\rangle$-congruent is strongly $\langle b,c\rangle$-congruent. So defining *$R$-congruence* as $\langle b,c\rangle$-congruence for all $R$-equivalent $\langle b,c\rangle$, and similarly for ***strong** $R$-congruence*, we have:

**(#):** Any pair of $R$-congruent valuations that occurs as a Kripke fixed point anywhere in the overall construction is strongly $R$-congruent.

And using (#), we can easily show by going through the construction that *every pair of valuations in the construction is strongly $R$-congruent.*[42]

Given this, it is a routine matter to "contract the model by $R$-equivalence": to replace the model by a reduced model where $R$-equivalent abstracts are identified. The value of every parameterized sentence is unaffected by the contraction, so naivety is preserved.

It might seem desirable to weaken (VI) by strengthening the assumption that $\langle j,v\rangle$ occur somewhere in the construction to the assumption that it be a recurring pair, in the sense that $v$ is recurring in the mid-level construction for $j$ and $j$ is either recurring in the revision-theoretic macro-construction or a stage along the way to the fixed point in the Brady. But the positive result couldn't hold for (I)-(V) plus this weakened version of (VI): Øgaard's proof rules it out. The argument for the positive result would fail since we could then prove only a weaker form of (#):

**($\#_w$):** Any pair of $R$-congruent valuations that occurs as a *recurring* Kripke fixed point in the overall construction is strongly $R$-congruent.

---

[42] Pick any $R$-congruent $j$. The mid-level construction from $j$ starts out from a valuation $v_{j,0}$ that assigns the same value to every $\rightarrow$-conditional, so it is trivially $R$-congruent; so by (#), $\langle j, v_{j,0}\rangle$ is strongly $R$-congruent. By the rules for the construction, this guarantees the $R$-congruence of the next member $v_{j,1}$ of the mid-level construction from $j$. Continuing in this way, we establish

**($):** Every $\rightarrow$-valuation $v_{j,\alpha}$ that occurs in the mid-level construction from a $R$-congruent $j$ is such that $\langle j, v_{j,\alpha}\rangle$ is strongly $R$-congruent.

Proceed analogously in the macro-construction: the starting $j_0$ assigns the same value to all $\rhd$-conditionals, so is trivially $R$-congruent, so ($) applies to it; and the rules of the macro-construction then guarantee that the next $j_1$ is $R$-congruent, and so on throughout the macro-construction.

And that doesn't suffice for the inductive proof of

**(%):** Every recurring pair of valuations in the construction is strongly $R$-congruent.

For in that proof (sketched in note 42), it was essential to start the induction from a pair $\langle j, v\rangle$ whose components are obviously $R$-congruent.

The upshot is that laws like $(\top \to \bot) \to \bot$, though in some sense valid, don't have the kind of "uniform validity" that is required for predicates coextensive by virtue of them to be sensibly regarded as expressing the same property. I don't see that that should be terribly upsetting.

## 7. Conclusion

I began by discussing the advantages of a naive theory of properties and propositions, and the paper has looked at several issues for such a theory, including (i) how it treats restricted quantification (and a conditional $\to$ used to define it), (ii) how it treats an ordinary conditional $\triangleright$, (iii) how it makes these interact so as to achieve the laws of restricted quantification we might expect, and (iv) what kinds of identity conditions for properties it permits.

A novelty in my treatment of (i) is the use of a continuum valued framework which generalizes Łukasiewicz continuum-valued semantics, and allows paradoxes that are treated there to be treated in essentially the same way while also providing a natural generalization that handles the paradoxes that Łukasiewicz semantics can't handle.

Under (ii) I argued that several methods do well: once one has decided how to treat an ordinary conditional without consideration of the paradoxes, any of a number of methods can be used to generalize it to handle the paradoxes peculiar to that conditional. This includes a simple method of Ross Brady's, provided it is given a novel tweak.

Under (iii) I argued that reasonable restricted quantifier laws don't depend too much on the details of the ordinary conditional used in stating the laws, they are largely settled by the laws of the $\to$ together with the "fiber bundle architecture" by which the treatment of the two conditionals is combined.

Under (iv) I adapted work in a joint paper with Harvey Lederman and Tore Fjetland Øgaard to show that there are some limitations on how coarse-grained one can take properties to be in a naive theory, but also to show that those limitations aren't that severe and that there are easy ways to attain fairly coarse-grained properties.

## References

[1] Beall, JC, Ross Brady, Allen Hazen, Graham Priest and Greg Restall 2006. "Relevant Restricted Quantification." *Journal of Philosophical Logic* 54: 87-103.

[2] Brady, Ross 1983. "The Simple Consistency of a Set Theory Based on the Logic CSQ". *Notre Dame Journal of Formal Logic* 24: 431-49.

[3] Burgess, John 1981. "Quick Completeness Proofs for Some Logics of Conditionals". *Notre Dame Journal of Formal Logic* 22: 76-84.

[4] Field, Hartry 2008. *Saving Truth from Paradox.*

[5] ———— 2010. "Reply to Greg Restall". *Philosophical Studies* 147: 460-67.

[6] ———— 2014. "Naive Truth and Restricted Quantification". *Review of Symbolic Logic* 7: 147-91.

[7] ———— 2016. "Indicative Conditionals, Restricted Quantification and Naive Truth". *Review of Symbolic Logic* 9: 181-208.

[8] Field, Hartry, Harvey Lederman and Tore Fjetland Øgaard forthcoming. "Prospects for a Naive Theory of Classes". *Notre Dame Journal of Formal Logic*.
[9] Gupta, Anil, and Nuel Belnap 1993. *The Revision Theory of Truth*. MIT Press.
[10] Hajek, Petr, Jeff Paris and John Sheperdson 2000. "The Liar Paradox and Fuzzy Logic", *Journal of Symbolic Logic* 65: 339-346.
[11] Kripke, Saul 1975. "Outline of a Theory of Truth." *Journal of Philosophy* 72: 690-716.
[12] Lewis, David 1981. "Ordering Semantics and Premise Semantics for Counterfactuals". *Journal of Philosophical Logic* 10: 217-34.
[13] Priest, Graham 2006. *In Contradiction*, second edition. Oxford University Press.
[14] ———————— 2008. *Introduction to Nonclassical Logic*, second edition. Cambridge University Press.
[15] Restall, Greg 1992. "Arithmetic and Truth in Łukasiewicz's Infinitely Valued Logic". *Logique et Analyse* 139-40: 303-12.
[16] ——————- 2010. "What are we to accept, and what are we to reject, in saving truth from paradox?" *Philosophical Studies* 147: 433-43.
[17] Rossi, Lorenzo forthcoming. "Graphs, Truth, and Conditionals." *Review of Symbolic Logic.*
[18] Schechter, Eric 2005. *Classical and Nonclassical Logics*. Princeton University Press.
[19] Stalnaker, Robert 1968. "A Theory of Conditionals." In N. Rescher, ed., *Studies in Logical Theory*, Blackwell Publishers.
[20] Standefer, Shawn 2015. "On Artifacts and Truth-preservation". *Australasian Journal of Logic* 12(3): 135-58.
[21] Visser, Albert 1984. "Four-Valued Semantics and the Liar". *Journal of Philosophical Logic* 13: 181-212.
[22] Yablo, Stephen 2003. "New Grounds for Naive Truth Theory". In JC Beall, ed., Liars and Heaps, *Oxford University Press*.