Journal of Philosophical Logic
https://doi.org/10.1007/s10992-021-09615-7

One Step is Enough

David Ripley[1]

Received: 21 June 2020 / Accepted: 19 June 2021 / Published online: 29 July 2021

Abstract
The recent development and exploration of mixed metainferential logics is a breakthrough in our understanding of nontransitive and nonreflexive logics. Moreover, this exploration poses a new challenge to theorists like me, who have appealed to similarities to classical logic in defending the logic ST, since some mixed metainferential logics seem to bear even more similarities to classical logic than ST does. There is a whole ST-based hierarchy, of which ST itself is only the first step, that seems to become more and more classical at each level. I think this seeming is misleading: for certain purposes, anyhow, metainferential hierarchies give us no reason to move on from ST. ST is indeed only the first step on a grand metainferential adventure; but one step is enough. This paper aims to explain and defend that claim. Along the way, I take the opportunity also to develop some formal tools and results for thinking about metainferential logics more generally.

Keywords Metainferences · Paradox · Nontransitive

The recent development and exploration of mixed metainferential logics (for example in [2, 3, 34, 35, 43]) is a breakthrough in our understanding of nontransitive and nonreflexive logics. Moreover, this exploration poses a new challenge to theorists like me, who have appealed to similarities to classical logic in defending the logic ST (for example in [69]), since some mixed metainferential logics seem to bear even more similarities to classical logic than ST does. There is a whole ST-based hierarchy, of which ST itself is only the first step, that seems to become more and more classical at each level. I think this seeming is misleading: for certain purposes, anyhow, metainferential hierarchies give us no reason to move on from ST. ST is indeed only the first step on a grand metainferential adventure; but one step is enough. This

✉ David Ripley
davewripley@gmail.com

[1] Philosophy Department SOPHIS Building 11, Monash University, 3800 Clayton VIC Australia

paper aims to explain and defend that claim. Along the way, I take the opportunity also to develop some formal tools and results for thinking about metainferential logics more generally.

In Section 1, I present the logic ST, and give a brief picture of some usual motivations for it. Section 2 lays out a picture of metainferences at all levels, and proves some general results about metainferential logics. Then, Section 3 returns to ST, now in metainferential-hierarchy form, and presents the ST hierarchy of [35]. Finally, Section 4 turns to the objection from [35, 43] and attempts to answer it.

1 ST

1.1 Language

Where not otherwise specified, I'm working with a propositional language $\mathcal{L}$ built on countably many atomic sentences, using the connectives $\neg$, $\wedge$, $\vee$, $\top$, and $\bot$, of arities 1, 2, 2, 0, 0 respectively, and intended as negation, conjunction, disjunction, verum, and falsum respectively.[1] For now, I use $\phi, \psi, \theta, \dots$ for arbitrary sentences, and $\Gamma, \Delta, \Sigma, \dots$ for sets of sentences. This is a simple and familiar language, and it is enough for present purposes.

1.2 Models

Models are usual three-valued valuations on the strong Kleene scheme (see for example [4] Ch. 8; [38] Ch. 7):

Definition 1 The *value algebra* $\mathfrak{V}$ is the set $\{1, .5, 0\}$ equipped with the following operations $\neg, \wedge, \vee, \top, \bot$, of arities 1, 2, 2, 0, 0, respectively:

$$\begin{aligned} \neg x &\text{ is } 1 - x \\ x \wedge y &\text{ is } \min(x, y) \\ x \vee y &\text{ is } \max(x, y) \\ \top &\text{ is } 1 \\ \bot &\text{ is } 0 \end{aligned}$$

A *model* is a homomorphism $\mathcal{L} \to \mathfrak{V}$. The set of all models is $\mathfrak{M}$.

I'll also have use for the subset of models that don't use the value .5:

Definition 2 A *2-valued model* is a model $\mathfrak{m}$ such that for every $\phi \in \mathcal{L}$, $\mathfrak{m}(\phi) \in \{1, 0\}$. The set of 2-valued models is $\mathfrak{M}_2$.

[1] I think of this language as an algebra, as in [14, §4.3, 4.4], and do not distinguish between the algebra itself and its carrier set, using the symbol $\mathcal{L}$ for both and trusting in context and the reader's generosity.

By combining these definitions, you can see that $\mathfrak{M}_2$ is the set of ordinary 2-valued Boolean valuations; it's handy to know that these are all in $\mathfrak{M}$.

1.3 Consequence Via Counterexamples

We can use these models to specify consequence relations by setting up a counterexample relation. I'll be looking at a few different consequence relations, built from different counterexample relations, all dealing with the same language $\mathcal{L}$ and the sets of models $\mathfrak{M}$ and $\mathfrak{M}_2$. So it'll be handy to have some general definitions.

Definition 3 An *inference* $[\Gamma \succ \Delta]$ is a pair of sets of sentences: *premises* Γ and *conclusions* Δ.[2] A *consequence relation* is a set of inferences; an inference is *valid according to* a consequence relation iff it is a member of the consequence relation.

Note that the definition of 'consequence relation' here is very general, imposing no conditions at all on such a set.

Definition 4 A *counterexample relation* is a relation between models and inferences. Given a counterexample relation X, a model $\mathfrak{m}$, and an inference $[\Gamma \succ \Delta]$, I write $\mathfrak{m}[\![\mathsf{X}]\!][\Gamma \succ \Delta]$ to mean that X relates $\mathfrak{m}$ to $[\Gamma \succ \Delta]$. Given a counterexample relation X, an inference $[\Gamma \succ \Delta]$ is X-*valid* (also written $\vDash^{\mathsf{X}} [\Gamma \succ \Delta]$) iff there is no model $\mathfrak{m}$ such that $\mathfrak{m}[\![\mathsf{X}]\!][\Gamma \succ \Delta]$.

This much works even for very strange choices of counterexample relation; and much of this paper develops tools for working with arbitrary counterexample relations. But it's sometimes nice to restrict attention to 'mixed' counterexample relations:[3]

Definition 5 Given $\mathsf{P}, \mathsf{C} \subseteq \{1, .5, 0\}$, the counterexample relation PC relates a model $\mathfrak{m}$ to an inference $[\Gamma \succ \Delta]$ iff $\mathfrak{m}[\Gamma] \subseteq \mathsf{P}$ and $\mathfrak{m}[\Delta] \cap \mathsf{C} = \emptyset$.[4]

[2] I will use usual sequent-style abbreviations without further comment, so for example $[\Gamma, \Gamma', \phi \succ]$ is the inference $[\Gamma \cup \Gamma' \cup \{\phi\} \succ \emptyset]$. For discussion of SET-SET and other frameworks, see for example [23, §1.21]; [47, Chs. 1, 2].

[3] Mixed consequence relations (that is, consequence relations based on mixed counterexample relations) have been studied in many forms. For early sources, see for example [16, 22, 29] and [17, Ch. 3], which itself points to [44, 52]. There are definite connections to 'Strawson-entailment', so-named in [57] after [51, pp. 68–69, 176–177] (see [46] for more on Strawson entailment, and see footnote 1 in [11] for more on the connection to mixed consequence). Some more recent applications involve particular attention to theories of vagueness (for example [63]; [50, §5.2]; [72]) or semantic properties such as truth and validity (for example [1, 32, 33, 64, 68]). More recently still, there has been an explosion of work on general frameworks for understanding such consequence relations: see [7, 9–11, 61, 70, 71].

[4] By $\mathfrak{m}[\Gamma]$, I mean $\{\mathfrak{m}(\gamma) | \gamma \in \Gamma\}$.

If you are comfortable with the idea of a *designated value*,[5] you can think of this as a slight generalization: we have a set P of *premise-designated* values and a possibly distinct set C of *conclusion-designated* values. Then we count a model as a counterexample to an inference iff the model premise-designates all the premises of the inference without conclusion-designating any of the conclusions.

I'll focus on two particular choices for these sets of designated values: the set $\mathsf{S} = \{1\}$ and the set $\mathsf{T} = \{1, .5\}$.[6] With just these two sets of values, we get four mixed consequence relations for this language and these models: $\vDash^{\mathsf{SS}}, \vDash^{\mathsf{TT}}, \vDash^{\mathsf{TS}}, \vDash^{\mathsf{ST}}$. The first two of these don't need the idea of mixed consequence, since they use the same standard of designation for premises and conclusions; these two are better-known as the consequence relations of the logics K3 and LP, respectively.[7] However, $\vDash^{\mathsf{TS}}$ and $\vDash^{\mathsf{ST}}$ really do make use of the flexibility of mixed consequence.

1.4 2-Valued Counterexamples

Let's restrict our attention for a moment to 2-valued models. The following Definition 6 is used only within this subsection:

Definition 6 Given a counterexample relation X, an inference $[\Gamma \succ \Delta]$ is X-2-*valid* (also written $\vDash^{\mathsf{X}}_2 [\Gamma \succ \Delta]$) iff there is no 2-valued model $\mathfrak{m}$ such that $\mathfrak{m}[\![\mathsf{X}]\!][\Gamma \succ \Delta]$.

Since 2-valued models never use the value .5, the difference between S and T disappears for these models. And so $\vDash^{\mathsf{SS}}_2, \vDash^{\mathsf{TT}}_2, \vDash^{\mathsf{TS}}_2$, and $\vDash^{\mathsf{ST}}_2$ are all the exact same consequence relation.

The consequence relation they all are is the usual consequence relation of classical propositional logic. This is because 2-valued models are usual Boolean valuations, and a 2-valued counterexample (in any of these four senses) to an inference $[\Gamma \succ \Delta]$ is a model that takes all sentences in Γ to the value 1 and all sentences in Δ to the value 0, which is the usual notion of a classical counterexample.

In what follows, it will be useful to be able to appeal to these 2-valued classical counterexamples. But rather than doing so by way of the mixed counterexample relations SS, TT, TS, ST, and in so doing needing to have a special definition of validity, as in Definition 6, I will rather introduce a fifth counterexample relation—one that is not a mixed counterexample relation, but is rather just the ordinary notion of a Boolean counterexample.

Definition 7 The counterexample relation CL relates a model $\mathfrak{m}$ to an inference $[\Gamma \succ \Delta]$ iff: $\mathfrak{m}$ is 2-valued, $\mathfrak{m}[\Gamma] \subseteq \{1\}$, and $\mathfrak{m}[\Delta] \subseteq \{0\}$.

[5] as explored for example in [14, Ch. 7] or [62, Ch. 3]

[6] These letters are for 'strict' and 'tolerant', and are taken from [63], where they are used ever so slightly differently.

[7] See again [4, Ch. 8]; [38, Ch. 7] for these.

CL is not a mixed counterexample relation: to determine whether a model is a CL counterexample to an inference, we need to know more than just the values the model assigns to sentences that occur in the inference. We need to know in addition whether the model is a 2-valued model, and this involves knowing about the values it assigns to every sentence in $\mathcal{L}$, not just those sentences that actually occur in the inference.

When a model is indeed a 2-valued model, it is a CL counterexample to an inference iff it is an SS counterexample iff it is a TT counterexample iff it is a TS counterexample iff it is an ST counterexample to that inference. So the CL counterexample relation captures restriction to 2-valued models, and it also captures the agreement of our four mixed counterexample relations over that restricted set of models.

1.5 Upshots

In this subsection, I record a few ideas it'll be handy to know as we go forward.

1.5.1 CL, ST, Vagueness, Truth

One important thing to note about $\vDash^{\mathsf{ST}}$ is that it is exactly the same consequence relation as $\vDash^{\mathsf{CL}}$. That is, $\vDash^{\mathsf{ST}} [\Gamma \succ \Delta]$ iff $\vDash^{\mathsf{CL}} [\Gamma \succ \Delta]$. Although ST and CL are two very different counterexample relations, these two counterexample relations determine the same consequence relation.

It is this combination of sameness and difference that underlies applications of ST to paradoxes of vagueness and truth.[8] Most of this paper leaves these applications to one side, but it's worth briefly introducing them here, to make clearer what some of the motivations are for exploring ST and its relatives.

In formulating logical theories of both vagueness and truth, the strong Kleene models I'm using here, or corresponding models for richer languages, have often proved fruitful.[9] However, these models are often associated with consequence relations weaker than $\vDash^{\mathsf{CL}}$,[10] and this has served as the basis for a number of objections to the resulting theories.[11]

However, by using ST counterexamples, these objections can be avoided. Theories of vagueness and truth based on strong Kleene models proceed by imposing certain restrictions on these models, corresponding to intuitive constraints such as tolerance in the case of vagueness or transparency in the case of truth. But since $\vDash^{\mathsf{ST}}$ validates every CL-valid inference while taking account of all strong Kleene models, these restricted classes of models only validate *more* inferences; they continue to validate all CL-valid inferences. So the resulting consequence relations for vagueness and truth come out to be *stronger* than $\vDash^{\mathsf{CL}}$. Any objections to the use of strong Kleene

[8] See for example [63, 66, 68].

[9] For example in [20, Ch. 11]; [21]; [25, Ch. III]; [26, 39, 45, 53, 55, 67].

[10] There are three usual candidates: the two here called $\vDash^{\mathsf{SS}}$ and $\vDash^{\mathsf{TT}}$, and the intersection of these two. See [6] for in-depth exploration of this intersection (in a different language), and [10] for discussion of intersections of mixed consequence relations more generally.

[11] For example in [19, Ch. 6], [24, pp. 103ff], [54], [60, p. 109ff].

models that are based on invalidating classically-valid inferences do not apply to ST.[12]

1.5.2 Transparent Truth and Cut

Let's look in just a tiny bit more detail at how this strategy can be used to offer a consequence relation for transparent truth. For this subsubsection, I'm officially talking about a first-order language with a truth predicate T and a special term $\langle\phi\rangle$ for each formula ϕ, although the details of this don't really matter.[13] A consequence relation for this language includes *transparent truth* iff ϕ and $T\langle\phi\rangle$ are interchangeable everywhere, even as subformulas of some other formula, without affecting the validity of any inference. A strong Kleene model for this language is a *transparent model* iff ϕ and $T\langle\phi\rangle$ are always assigned the same value by the model. One way to give consequence relations with transparent truth is by restricting attention to transparent models.

Given certain assumptions, in such languages we can often form paradoxical sentences, for example a liar sentence λ that is $\neg T\langle\lambda\rangle$. If we have such a sentence, however,then there can be no transparent 2-valued models.[14] If a model assigns the value 1 to λ, then by transparency it would have to assign 1 to $T\langle\lambda\rangle$ and so 0 to $\neg T\langle\lambda\rangle$, which is just λ again; so this doesn't work. If a model assigns the value 0 to λ, then by transparency it would have to assign 0 to $T\langle\lambda\rangle$ and so 1 to $\neg T\langle\lambda\rangle$, which is just λ again; so this doesn't work either. But if the model is 2-valued, there is no other value for λ to receive. So if we restrict our attention to the set of transparent 2-valued models, we are restricting our attention to the empty set—and no models means no counterexamples, so every inference comes out valid. This total consequence relation does indeed feature transparent truth, but that's about the only nice thing you can say about it.

On the other hand, even in the presence of paradoxical sentences of all kinds, there are many transparent models; it's just that none of them are 2-valued. Indeed, for any model $\mathfrak{m}$ at all, there are transparent models that differ from $\mathfrak{m}$ only in their handling of the truth predicate.[15] Any mixed consequence relation over these transparent models includes transparent truth. Let's focus for a moment on the following notion:

---

[12] This is roughly the dual of the central argument of [66]. There, I defend classical logic against objections from tolerant vagueness and transparent truth, proceeding proof-theoretically. Here, I'm sketching a way to defend tolerant vagueness and transparent truth against objections from classical logic, proceeding model-theoretically.

[13] There are a number of ways to achieve such naming devices; see [13, 26, 64] for three different options that work in this kind of setting.

[14] See [18] for ways to achieve transparent 2-valued models by avoiding such sentences.

[15] For demonstration that such models exist, see for example [27, 30, 45], all of which demonstrate this fact in similar ways. [45, p. 4] includes an interesting historical paragraph (although frustratingly without citations), concluding: "The problem of priority is not very acute: Fitch started in the mid-thirties and gets the main credit". Other useful discussions include [15, Chs. 3, 4, 16]; [28, 56]. Note also [8, 48, 49], which are closely related.

Definition 8 A model is an STT counterexample to an inference iff: it is transparent and it is an ST counterexample to that inference.

Since all STT counterexamples to an inference are also ST counterexamples to that inference, $\models^{\mathsf{STT}}$ validates every inference that $\models^{\mathsf{ST}}$ does. That is, $\models^{\mathsf{STT}}$ validates every classically-valid inference.[16] However, as it is built on transparent models, $\models^{\mathsf{STT}}$ also includes a transparent truth predicate, even in the presence of paradoxical sentences of all kinds. So $\models^{\mathsf{STT}}$ occupies an interesting position among logical approaches to truth: it is nonclassical enough in its models to achieve transparency, but it does this without invalidating any classically valid inferences.

But the pressure generated by the paradoxes is always vented somewhere. In this case, it's here: $\models^{\mathsf{STT}}$ is not closed under the rule of cut. That is, there are Γ, Δ, and ϕ with $\models^{\mathsf{STT}} [\Gamma \succ \Delta, \phi]$ and $\models^{\mathsf{STT}} [\phi, \Gamma \succ \Delta]$, but with $\not\models^{\mathsf{STT}} [\Gamma \succ \Delta]$. In particular, let λ be a liar sentence as above, and let $\Gamma = \Delta = \emptyset$. As we saw, λ must receive the value .5 in any transparent model. So there is no STT counterexample to either $[\succ\lambda]$ or $[\lambda\succ]$, and thus $\models^{\mathsf{STT}} [\succ\lambda]$ and $\models^{\mathsf{STT}} [\lambda\succ]$. But every transparent model is an STT counterexample to $[\succ]$; and as there are such models, $\not\models^{\mathsf{STT}} [\succ]$.

How can $\models^{\mathsf{STT}}$ maintain the validity of all classically valid inferences while not being closed under such a classical-seeming principle as cut? Well, because instances of cut are not inferences. They are a different kind of thing: metainferences. The failure of $\models^{\mathsf{STT}}$ to be closed under cut is one motivation, then, for studying metainferences.

2 Higher Levels

For this section, you can forget entirely about the details of our language $\mathcal{L}$ and our models $\mathfrak{M}$. I will lay out one particular way of thinking of metainferences, metametainferences, and so on, but I will do this assuming only that we have some set of sentences and some set of models. In the next section, we'll have use again for the propositional language $\mathcal{L}$ and the strong Kleene models $\mathfrak{M}$, but there is a bunch of structure worth seeing that does not depend at all on which language or which models we have in mind, and that structure is the topic of this section. This material follows [3, 35, 43] in many places. But I hope the extra generality provides a useful perspective.[17]

2.1 Meta^{n}inferences, meta^{n}Counterexamples, meta^{n}Consequence

The *levels* are $\ell = \{-1\} \cup \mathbb{N}$; throughout, variables i, j, k, m, n range over levels. From here forward, I will use $\phi, \psi, \ldots$ for minferences of all levels, and $\Gamma, \Delta, \ldots$

[16] Above, I pointed out that $\models^{\mathsf{ST}}$ matches $\models^{\mathsf{CL}}$ in our propositional language; that remains true in this first-order language. See for example [64, 65].

[17] This section consists, largely unavoidably I'm afraid, of a flurry of definitions and very simple proofs of facts interrelating them. There is as yet no standard way of approaching this material, so a lot still has to be done from scratch.

for sets of minferences. (That means I'll be using $\Gamma, \Delta, \ldots$ for meta^{n}consequence relations and full consequence relations as well as for the components of minferences, since sets of minferences play both roles.)

Definition 9 A *meta^{-1}inference* is a sentence. A *meta^{n+1}inference* $[\Gamma \succ \Delta]$ is a pair of sets of meta^{n}inferences: *premise meta^{n}inferences* Γ and *conclusion meta^{n}inferences* Δ. $\mathbf{M}_n$ is the set of all meta^{n}inferences, and $\mathbf{M}$ is $\bigcup_{n \in \ell} \mathbf{M}_n$. The members of $\mathbf{M}$ are the *minferences*.[18]

Definition 9 has the consequence that a meta^{0}inference is an inference.[19] For examples at higher levels, $[[\succ \phi] \succ [\phi \succ \phi, \psi], [\psi \succ \psi]]$ is a meta^{1}inference (with one premise inference and two conclusion inferences),

$$[[[\succ \phi] \succ [\phi \succ \phi, \psi], [\psi \succ \psi]], [[\psi \succ \psi], [\phi \succ] \succ [\succ \psi]] \succ [[\phi, \psi \succ \phi, \psi] \succ [\phi \succ \psi]]]$$

is a meta^{2}inference (with two premise meta^{1}inferences and one conclusion meta^{1}inference), and so on. That last example is awful to parse, I know. It wasn't that easy to type, either. Fortunately, nothing in this paper is actually going to require you to parse things like that, so we're ok.

Definition 10 A *meta^{n}consequence relation* is a subset of $\mathbf{M}_n$, and a *full consequence relation* is a subset of $\mathbf{M}$. Given a full consequence relation Σ and a level n, the meta^{n}consequence relation $\Sigma(n)$ is $\Sigma \cap \mathbf{M}_n$. All of these are *consequence relations*.

Definition 11 A *meta^{n}counterexample relation* is a relation between models and $\mathbf{M}_n$. A *full counterexample relation* is a relation between models and $\mathbf{M}$. Given a full counterexample relation X and a level n, the meta^{n}counterexample relation $\mathsf{X}(n)$ is the restriction of X in its codomain to $\mathbf{M}_n$.[20] All of these are *counterexample relations*. Given a counterexample relation X, a model $\mathfrak{m}$, and a minference μ, I write $\mathfrak{m}[\![\mathsf{X}]\!]\mu$ to mean that X relates $\mathfrak{m}$ to μ.

[18] 'Metainference' appears to be a relatively settled name for what I'm here calling 'meta^{1}inferences', so I needed a different name for the general collection. Note also that this definition joins [43] in extending the SET-SET framework to all levels, in contrast with [3, 35], which restrict their attention to meta^{n+1}inferences with a single conclusion meta^{n}inference when $n \geq 0$.

[19] Even the small existing literature here is already not uniform in its level numbering. My level numbering here agrees with [35], and differs from [3, 43]. The numbering I'm using has the nice consequence that meta^{1}inferences are what are elsewhere called 'metainferences', meta^{2}inferences are 'metametainferences', etc.

[20] That is, I will treat a full counterexample relation indifferently as on the one hand a relation between models and $\mathbf{M}$ and on the other a function with *dependent type* $(n : \ell) \to \mathbf{M}_n$. For more on dependent types, see for example [31, 59].

It's important to keep Definitions 10 and 11 clearly distinguished.[21] This is because much of the work to be done focuses on counterexample relations, but the applications to be made of this work involve mainly consequence relations. The ideas are not totally separate, however; they do relate to each other in a useful way. Given a counterexample relation and a collection of models, we can determine a consequence relation. Since I'm holding our collection of models fixed, we can talk simply of the consequence relation determined by a given counterexample relation:

Definition 12 Given a meta^{n}counterexample relation X, the meta^{n}consequence relation $\mathcal{C}(\mathsf{X})$ is the set $\{\mu \in \mathbf{M}_n | \text{there is no model } \mathfrak{m} \text{ with } \mathfrak{m}[\![\mathsf{X}]\!]\mu\}$. Given a full counterexample relation X, the full consequence relation $\mathcal{C}(\mathsf{X})$ is the set $\{\mu \in \mathbf{M} | \text{there is no model } \mathfrak{m} \text{ with } \mathfrak{m}[\![\mathsf{X}]\!]\mu\} = \bigcup_{n \in \ell} \mathcal{C}(\mathsf{X}(n))$.

This determination is not invertible. Multiple distinct meta^{n}counterexample relations can determine the same meta^{n}consequence relation, and multiple distinct full counterexample relations can determine the same full consequence relation. (We've already met an example of the first sort: ST and CL are distinct meta^{0}counterexample relations, and as we've already seen, $\mathcal{C}(\mathsf{ST}) = \mathcal{C}(\mathsf{CL})$. Later, we will come to the best-known example of the second sort.) This is one reason it is important to stay focussed mainly on counterexample relations rather than consequence relations. Consequence relations simply are not carrying enough information on their own; the extra fineness of grain provided by counterexample relations plays an important role.

2.2 Lifting and Lowering

It will be useful to explore relations between levels. The following operations of *lifting* and *lowering* on counterexample relations come in for heavy use:

Definition 13 (Lifting) Given a meta^{n}counterexample relation X, its lifting $\uparrow\mathsf{X}$ is the meta^{n+1}counterexample relation such that for any model $\mathfrak{m}$ and any meta^{n+1}inference $[\Gamma \succ \Delta]$, we have $\mathfrak{m}[\![\uparrow\mathsf{X}]\!][\Gamma \succ \Delta]$ iff: $\mathfrak{m}[\![\mathsf{X}]\!]\delta$ for all $\delta \in \Delta$ but there is no $\gamma \in \Gamma$ with $\mathfrak{m}[\![\mathsf{X}]\!]\gamma$.

Definition 14 (Lowering) Given a meta^{n+1}counterexample relation X, its lowering $\downarrow\mathsf{X}$ is the meta^{n}counterexample relation such that for any model $\mathfrak{m}$ and any meta^{n-1}inference μ, we have $\mathfrak{m}[\![\downarrow\mathsf{X}]\!]\mu$ iff $\mathfrak{m}[\![\mathsf{X}]\!][\succ\mu]$.

We also have that lowering is a retraction (aka a left inverse) of lifting:

[21] Pailos [35] speaks of 'consequence relations', but I'm not entirely sure in what sense. Full counterexample relations correspond most directly to [43]'s 'logics'. (My level numbering, however, agrees with [35]'s, and differs from [43]'s by one: in [43], sentences are at level 0, where as here, they are at −1.)

Fact 1 *For any meta^n counterexample relation* X*, we have* $\mathsf{X} = \downarrow\uparrow \mathsf{X}$.

Proof For any model $\mathfrak{m}$ and meta^{n}inference μ, we have $\mathfrak{m}[\![\downarrow\uparrow \mathsf{X}]\!]\mu$ iff $m[\![\uparrow \mathsf{X}]\!][\succ\mu]$. This obtains iff $\mathfrak{m}[\![\mathsf{X}]\!]\mu$ but there is no $\gamma \in \emptyset$ with $\mathfrak{m}[\![\mathsf{X}]\!]\gamma$, which is to say it obtains iff $\mathfrak{m}[\![\mathsf{X}]\!]\mu$. □

However, the lowering operation is not injective, and so it has no full inverse; in particular, it is not always the case that $\uparrow\downarrow \mathsf{X} = \mathsf{X}$. For example, consider ST. $\downarrow \mathsf{ST}$ relates a model $\mathfrak{m}$ to a meta^{-1}inference (which is to say a sentence) ϕ iff ST relates that model to $[\succ\phi]$; this holds iff $\mathfrak{m}(\phi) = 0$. And so $\uparrow\downarrow \mathsf{ST}$ relates a model $\mathfrak{m}$ to a meta^{0}inference $[\Gamma \succ \Delta]$ iff $\mathfrak{m}(\delta) = 0$ for all $\delta \in \Delta$ and $m(\gamma) \neq 0$ for all $\gamma \in \Gamma$. This is not ST but TT.[22]

2.2.1 Coherence

With these notions of lifting and lowering in hand, we can ask how the different levels of a full counterexample relation relate to each other. Two ideas will be needed: *downward coherence* and *n-upward coherence*.

Definition 15 A full counterexample relation X is *downward coherent* iff for all levels n, $\mathsf{X}(n) = \downarrow \mathsf{X}(n+1)$.[23]

Definition 16 A full counterexample relation X is *n-upward coherent* iff for all levels $m \geq n$, $\mathsf{X}(m+1) = \uparrow \mathsf{X}(m)$.

That is, a full counterexample relation is downward coherent iff whenever we consider its restriction to two adjacent levels, the lower level is the lowering of the upper one. And it is n-upward coherent iff whenever we consider its restriction to two adjacent levels of at least level n, the upper level is the lift of the lower one. The reason for the n in the notion of n-upward coherence, and its absence in the notion of downward coherence, is just that the particular full counterexample relations to be considered here are all downward coherent, but many are only n-upward coherent for certain n. (Some aren't n-upward coherent for any n.) But more on that when we get to it.

[22] Indeed, in general for a mixed counterexample relation PC, we have $\uparrow\downarrow \mathsf{PC} = \mathsf{CC}$, for the same reason. For the generalization of this to all levels, see Fact 5.

[23] Scambler [43, pp. 359–361] discusses 'the Eqs. 2–3 equivalence', which is closely related. But there is a difference: the Eqs. 2–3 equivalence applies to full consequence relations rather than full counterexample relations, and while a downward coherent full counterexample relation always determines a full consequence relation obeying the Eqs. 2–3 equivalence, there are non-downward coherent full counterexample relations that do so as well.

2.2.2 Determining Full Counterexample Relations

By Definition 11, any full counterexample relation determines a meta^{n}counterexample relation immediately, by restriction. With lifting and lowering in hand, we can go the other way, taking any meta^{n}counterexample relation to determine a full counterexample relation in a natural way as well:[24]

Definition 17 For $\circ \in \{\uparrow, \downarrow\}$, let $\circ^0 \mathsf{X} = \mathsf{X}$, and let $\circ^{i+1} \mathsf{X} = \circ^i \circ \mathsf{X}$. Then a meta^{n}counterexample relation X *determines* the full counterexample relation $\widehat{\mathsf{X}}$ as follows:

- for $k \geq n$, $\widehat{\mathsf{X}}(k) = \uparrow^{k-n} \mathsf{X}$, and
- for $k \leq n$, $\widehat{\mathsf{X}}(k) = \downarrow^{n-k} \mathsf{X}$.[25]

Fact 2 *Any full counterexample relation determined by a meta^{n}counterexample relation is downward coherent.*

Proof Take a $\widehat{\mathsf{X}}$ determined by some meta^{n}counterexample relation X. We need to show that for any m, $\widehat{\mathsf{X}}(m) = \downarrow(\widehat{\mathsf{X}}(m+1))$. There are two cases:

- $m \geq n$: In this case, $\widehat{\mathsf{X}}(m) = \uparrow^{m-n} \mathsf{X}$. By Fact 1, $\uparrow^{m-n} \mathsf{X} = \downarrow\uparrow\uparrow^{m-n} \mathsf{X}$. And $\downarrow\uparrow\uparrow^{m-n} \mathsf{X} = \downarrow\uparrow^{(m+1)-n} \mathsf{X} = \downarrow \widehat{\mathsf{X}}(m+1)$.
- $m < n$: In this case, $\widehat{\mathsf{X}}(m) = \downarrow^{n-m} \mathsf{X}$. Since $n - m \geq 1$, this is $\downarrow\downarrow^{n-(m+1)} \mathsf{X}$. And $\downarrow\downarrow^{n-(m+1)} \mathsf{X} = \downarrow \widehat{\mathsf{X}}(m+1)$. □

Fact 3 *Any full counterexample relation determined by a meta^{n}counterexample relation is n-upward coherent.*

Proof Take a full counterexample relation $\widehat{\mathsf{X}}$ determined by a meta^{n}counterexample relation X. We need to show that for any $m \geq n$, $\widehat{\mathsf{X}}(m+1) = \uparrow(\widehat{\mathsf{X}}(m))$. We have $\widehat{\mathsf{X}}(m+1) = \uparrow^{(m+1)-n} \mathsf{X} = \uparrow\uparrow^{m-n} \mathsf{X} = \uparrow \widehat{\mathsf{X}}(m)$. □

Fact 4 *For any n, if a full counterexample relation X is both downward coherent and n-upward coherent, then it is determined by its nth layer $\mathsf{X}(n)$.*

Proof We need to show for all m that $\mathsf{X}(m) = ([\widehat{\mathsf{X}(n)}](m))$. There are two cases:

- $m \geq n$: In this case, $([\widehat{\mathsf{X}(n)}](m)) = \uparrow^{m-n}(\mathsf{X}(n))$. Since X is n-upward coherent, this is $\mathsf{X}(m)$.
- $m < n$: In this case, $([\widehat{\mathsf{X}(n)}](m)) = \downarrow^{n-m}(\mathsf{X}(n))$. Since X is downward coherent, this is $\mathsf{X}(m)$. □

[24] Compare [43, Defn. 16], which is less general but aimed at what I take to be a similar idea.
[25] These clauses overlap when $k = n$, but in this case they also agree, so no harm done.

Corollary 1 *A full counterexample relation is determined by its nth layer iff it is downward coherent and n-upward coherent.*

Proof Summing up Facts 2, 3 and 4. □

Let a full counterexample relation be *principal* iff it is determined by some meta^{n}counterexample relation, for some n. Then I will only directly consider two non-principal full counterexample relation in this paper: what I will later introduce as ST_ω and TS_ω. All other full counterexample relations to be considered here are principal, and so downward coherent and n-upward coherent, at least for some n.

Note that a meta^{0}counterexample relation X, on its own, says nothing at all about validity of meta^{n}inferences for $n \neq 0$. Despite this, there is a tendency to move quickly from X to $\widehat{\mathsf{X}}$, at least for some purposes. For example, [43, p. 360, notation changed] says, "[A]bsent any other reasons for suspicion one should probably take $\widehat{\mathsf{X}}$ to be what someone has in mind if they only specify X". I don't think this tendency is warranted. Most of the time, when someone has specified a meta^{0}counterexample relation (which is to say an ordinary counterexample relation), they do not have the world of all higher minferences, full counterexample relations, etc, in mind at all. They are often just focussed on validity for meta^{0}inferences (which is to say inferences). In what follows, I will look at particular meta^{n}counterexample relations X in part via the full counterexample relations $\widehat{\mathsf{X}}$ they determine, and I'll specify particular principal full counterexample relations by meta^{n}counterexample relations that determine them. But I'll still try to be careful about the distinction.

2.2.3 Climbing by Slashing

The lift operation $\uparrow$ determines a meta^{n+1}counterexample relation from a meta^{n}counterexample relation by applying that meta^{n}counterexample relation uniformly to premise meta^{n}inferences and conclusion meta^{n}inferences. But, corresponding to Definition 5, we can also determine a meta^{n+1}counterexample relation from two possibly distinct meta^{n}counterexample relations, one for premise meta^{n}inferences and the other for conclusion meta^{n}inferences:

Definition 18 Given meta^{n}counterexample relations P and C, the meta^{n+1}counterexample relation P/C is determined as follows: for any model $\mathfrak{m}$ and meta^{n+1}inference $[\Gamma \succ \Delta]$, we have $\mathfrak{m}[\![\mathsf{P}/\mathsf{C}]\!][\Gamma \succ \Delta]$ iff: for all $\delta \in \Delta$ we have $\mathfrak{m}[\![\mathsf{C}]\!]\delta$ and there is no $\gamma \in \Gamma$ with $\mathfrak{m}[\![\mathsf{P}]\!]\gamma$.

Fact 5 *For any meta^{n}counterexample relations* $\mathsf{X}, \mathsf{P}, \mathsf{C}$, *we have* $\uparrow\mathsf{X} = \mathsf{X}/\mathsf{X}$, *and* $\downarrow(\mathsf{P}/\mathsf{C}) = \mathsf{C}$.

Proof Unraveling Definitions 13, 14 and 18 □

2.3 Moving to Consequence Relations

While some of the work to follow involves only counterexample relations, some also involves the consequence relations they determine, and it's useful to have a few tools here as well.

2.3.1 Lowering Consequence Relations

Lowering does not require the fine grain of counterexample relations; we can lower consequence relations directly, in a way that agrees with lowering for counterexample relations.

Definition 19 Given a meta^{n+1}consequence relation Σ, let $\downarrow\Sigma$ be the meta^{n}consequence relation $\{\mu \in \mathbf{M}_n | [\succ\mu] \in \Sigma\}$.

Fact 6 *For any meta^{n+1}counterexample relation* X *and meta^{n}inference μ, we have* $\mathcal{C}(\downarrow \mathsf{X}) = \downarrow \mathcal{C}(\mathsf{X})$.

Proof Combining Definitions 12, 14 and 19. □

However, the same is not possible for lifting. This is because there can be meta^{n}counterexample relations X and Y with $\mathcal{C}(\mathsf{X}) = \mathcal{C}(\mathsf{Y})$ but $\mathcal{C}(\uparrow \mathsf{X}) \neq \mathcal{C}(\uparrow \mathsf{Y})$. Indeed, we have already met such: $\mathcal{C}(\mathsf{ST}) = \mathcal{C}(\mathsf{CL})$, but $\mathcal{C}(\uparrow \mathsf{ST}) \neq \mathcal{C}(\uparrow \mathsf{CL})$.[26] So there cannot be any operation $\uparrow$ on consequence relations such that in general $\uparrow \mathcal{C}(\mathsf{X}) = \mathcal{C}(\uparrow \mathsf{X})$. Lifting depends on information carried by a counterexample relation that is lost in the move to the consequence relation it determines.

2.3.2 Agreement

I'll be occupied in what follows with the situation where two distinct full counterexample relations determine the same meta^{n}consequence relation, or when they determine the same full consequence relation.

Definition 20 Full counterexample relations X, Y *agree at level n* (written $\mathsf{X} \approx_n \mathsf{Y}$) iff $\mathcal{C}(\mathsf{X}(n)) = \mathcal{C}(\mathsf{Y}(n))$. They *agree fully* (written $\mathsf{X} \approx \mathsf{Y}$) iff $\mathcal{C}(\mathsf{X}) = \mathcal{C}(\mathsf{Y})$.

Fact 7 *If* X *and* Y *are downward coherent and* $\mathsf{X} \approx_n \mathsf{Y}$*, then* $\mathsf{X} \approx_m \mathsf{Y}$ *for all* $m \leq n$.

Proof It suffices to show that $\mathsf{X} \approx_{n-1} \mathsf{Y}$, assuming $n \geq 0$. Since both are downward coherent, $\mathsf{X}(n-1) = \downarrow \mathsf{X}(n)$ and $\mathsf{Y}(n-1) = \downarrow \mathsf{Y}(n)$. So $\mathcal{C}(\mathsf{X}(n-1)) = \mathcal{C}(\downarrow \mathsf{X}(n))$, which by Fact 6 is $\downarrow \mathcal{C}(\mathsf{X}(n))$, and similarly $\mathcal{C}(\mathsf{Y}(n-1)) = \downarrow \mathcal{C}(\mathsf{Y}(n))$. As $\mathcal{C}(\mathsf{X}(n)) = \mathcal{C}(\mathsf{Y}(n))$, then $\mathcal{C}(\mathsf{X}(n-1)) = \mathcal{C}(\mathsf{Y}(n-1))$. □

[26] This follows from Facts 15 and 9, to come.

So as long as we're dealing with downward coherent counterexample relations, agreement *at* level n suffices for agreement *up to and including* level n.

There is no corresponding fact involving n-upward coherence. Note the involvement of Fact 6 in the above proof, and recall that this has no analog for lifting instead of lowering. Indeed, $\widehat{\mathsf{ST}}$ and $\widehat{\mathsf{CL}}$ are 0-upward coherent (since determined by meta^{0}counterexample relations), and $\widehat{\mathsf{ST}} \approx_0 \widehat{\mathsf{CL}}$, but $\widehat{\mathsf{ST}} \not\approx_1 \widehat{\mathsf{CL}}$.[27]

3 The ST_n and TS_n Hierarchies

Enough generalities. Let's return now to our propositional language $\mathcal{L}$ and strong Kleene models $\mathfrak{M}$, and turn to the central idea of [35, 43]: the ST_n and TS_n hierarchies of counterexample relations. This proceeds by mixing at every level.

Definition 21

- - ST_{-1} relates a model $\mathfrak{m}$ to a meta^{-1}inference (sentence) ϕ iff $\mathfrak{m}(\phi) = 0$
  - TS_{-1} relates a model $\mathfrak{m}$ to a meta^{-1}inference (sentence) ϕ iff $\mathfrak{m}(\phi) \neq 1$
- - ST_{n+1} is $\mathsf{TS}_n/\mathsf{ST}_n$
  - TS_{n+1} is $\mathsf{ST}_n/\mathsf{TS}_n$

This captures ST and TS, since $\mathsf{ST}_0 = \mathsf{ST}$ and $\mathsf{TS}_0 = \mathsf{TS}$, but it goes much farther, giving ST-like and TS-like meta^{n}counterexample relations for every n.[28]

3.1 Exploring Both Hierarchies

Here, I present some fact about these hierarchies.

First, note that lowering brings us back down them:

Fact 8 $\mathsf{ST}_n = \downarrow \mathsf{ST}_{n+1}$ *and* $\mathsf{TS}_n = \downarrow \mathsf{TS}_{n+1}$.

Proof By Fact 5. □

But lifting does not bring us up; it does not even bring us to a place that agrees on consequence:

Fact 9 *For all* n, $\widehat{\mathsf{ST}_n} \not\approx_{n+1} \widehat{\mathsf{ST}_{n+1}}$ *and* $\widehat{\mathsf{TS}_n} \not\approx_{n+1} \widehat{\mathsf{TS}_{n+1}}$.[29]

[27] As in Footnote 26, this follows from Facts 15 and 9.

[28] Scambler [43] names the ST_n hierarchy after T, and the TS_n hierarchy after S, with some justice. I don't do that here mainly to try to limit the potential confusion from the fact that my level numbering is different. So each of my ST_ns is [43]'s T_{n+1}, which in turn corresponds to [35]'s $\boldsymbol{CM_n}$.

[29] See also [35, Thm. 5.5] and [43, Lemma 22].

Proof By Fact 5, $\uparrow \mathsf{ST}_n = \mathsf{ST}_n/\mathsf{ST}_n$ and $\uparrow \mathsf{TS}_n = \mathsf{TS}_n/\mathsf{TS}_n$. So for the first claim, it suffices to give a meta^{n+1}inference in $\mathcal{C}(\mathsf{TS}_n/\mathsf{ST}_n(n+1))$ but not in $\mathcal{C}(\mathsf{ST}_n/\mathsf{ST}_n(n+1))$; and for the second, it suffices to give a meta^{n+1}inference in $\mathcal{C}(\mathsf{TS}_n/\mathsf{TS}_n(n+1))$ but not in $\mathcal{C}(\mathsf{ST}_n/\mathsf{TS}_n(n+1))$.

Let ϕ be an atomic sentence. Let $\phi_{-1} = \phi$ and $\phi_{k+1} = [\succ \phi_k]$. Let $(\neg\phi)_{-1} = \neg\phi$ and $(\neg\phi)_{k+1} = [\succ(\neg\phi)_k]$. Note by induction on k that for all k: every model is a TS_k meta^{k}counterexample to at least one of ϕ_k, $(\neg\phi)_k$, and any model $\mathfrak{m}$ with $\mathfrak{m}(\phi) = .5$ is not an ST_k meta^{k}counterexample to either of ϕ_k, $(\neg\phi)_k$. Since ϕ is atomic, there are plenty of such models.

Now let $\mu_{k+1} = [\phi_k, (\neg\phi)_k \succ]$; by the above, μ_{k+1} is in both $\mathcal{C}(\mathsf{TS}_n/\mathsf{ST}_n(n+1))$ and $\mathcal{C}(\mathsf{TS}_n/\mathsf{TS}_n(n+1))$, but not in either of $\mathcal{C}(\mathsf{ST}_n/\mathsf{ST}_n(n+1))$ or $\mathcal{C}(\mathsf{ST}_n/\mathsf{TS}_n(n+1))$. □

This in turn assures us that all levels of these hierarchies are distinct:

Corollary 2 *if* $i < j$ *and* $i < k$, *then* $\widehat{\mathsf{ST}_i} \not\approx_j \widehat{\mathsf{ST}_k}$ *and* $\widehat{\mathsf{TS}_i} \not\approx_j \widehat{\mathsf{TS}_k}$.

Proof Suppose $\widehat{\mathsf{ST}_i} \approx_j \widehat{\mathsf{ST}_k}$. By Fact 2 both $\widehat{\mathsf{ST}_i}$ and $\widehat{\mathsf{ST}_k}$ are downward coherent, and so by Fact 7, $\widehat{\mathsf{ST}_i} \approx_{i+1} \widehat{\mathsf{ST}_k}$. By Fact 5, $\widehat{\mathsf{ST}_k}(i+1) = \mathsf{ST}_{i+1}$, so $\widehat{\mathsf{ST}_i} \approx_{i+1} \widehat{\mathsf{ST}_{i+1}}$. This contradicts Fact 9. The same argument works on the TS side of the hierarchy. □

In addition to the members of these hierarchies, we can use them to specify two additional full counterexample relations.

Definition 22 ST_ω is the full counterexample relation such that $\mathsf{ST}_\omega(n) = \mathsf{ST}_n$. TS_ω is the full counterexample relation such that $\mathsf{TS}_\omega(n) = \mathsf{TS}_n$.

Unlike all the other full counterexample relations in this paper, these two full counterexample relations are *not* principal; they are not determined by any meta^{n}counterexample relation, for any n. If they were, by Fact 3, they would be n-upward coherent, but they are not:

Fact 10 ST_ω *and* TS_ω *are downward coherent, and they are not* n*-upward coherent for any* n.

Proof Downward coherence is from Fact 5 and Definition 21. To see that they are not n-upward coherent, choose any n and suppose ST_ω to be n-upward coherent. This implies that $\uparrow \mathsf{ST}_n = \mathsf{ST}_{n+1}$, which contradicts Fact 9. □

To explore these hierarchies more fully, I'll use the following ordering on strong Kleene models:

Definition 23 Let $\sqsubseteq$ be the smallest partial order on $\{1, .5, 0\}$ such that $.5 \sqsubseteq 1$ and $.5 \sqsubseteq 0$. For models $\mathfrak{m}, \mathfrak{m}'$, let $\mathfrak{m} \sqsubseteq \mathfrak{m}'$ iff for all sentences ϕ, we have $\mathfrak{m}(\phi) \sqsubseteq \mathfrak{m}'(\phi)$.

Fact 11 *If* $\mathfrak{m}(\phi) \sqsubseteq \mathfrak{m}'(\phi)$ *for all atomic sentences* ϕ, *then* $\mathfrak{m} \sqsubseteq \mathfrak{m}'$.

Proof By induction on sentence formation, noting that all connectives are monotonic wrt $\sqsubseteq$; that is, for any $u, u', v, v' \in \{1, .5, 0\}$:

- if $u \sqsubseteq u'$, then $\neg u \sqsubseteq \neg u'$;
- if $u \sqsubseteq u'$ and $v \sqsubseteq v'$, then $u \wedge v \sqsubseteq u' \wedge v'$; and
- the same goes, mutatis mutandis, for $\vee, \top, \bot$. □

Fact 12 *If* $\mathfrak{m} \sqsubseteq \mathfrak{m}'$*, then: if* $\mathfrak{m}$ *is an* ST_n *meta*^{n}*counterexample to a meta*^{n}*inference* μ*, then so is* $\mathfrak{m}'$*; and if* $\mathfrak{m}'$ *is a* TS_n *meta*^{n}*counterexample to a meta*^{n}*inference* μ*, then so is* $\mathfrak{m}$*.*[30]

Proof By induction on n. □

Corollary 3 *If* $\mathfrak{m} \sqsubseteq \mathfrak{m}'$*, then: if* $\mathfrak{m}$ *is an* ST_ω *counterexample to a minference* μ*, then so is* $\mathfrak{m}'$*; and if* $\mathfrak{m}'$ *is a* TS_ω *counterexample to a minference* μ*, then so is* $\mathfrak{m}$*.*

Finally, note the relation to CL and 2-valued models. We know that a 2-valued model is a CL meta^{0}counterexample to a meta^{0}inference μ iff it is an ST meta^{0}counterexample to μ iff it is a TS meta^{0}counterexample to μ. This relationship extends to all levels:

Fact 13 *Take any 2-valued model* $\mathfrak{m}$ *and any meta*^{n}*inference* μ*. Then* $\mathfrak{m}[\![\widehat{\mathsf{CL}}(n)]\!]\mu$ *iff* $\mathfrak{m}[\![\mathsf{ST}_n]\!]\mu$ *iff* $\mathfrak{m}[\![\mathsf{TS}_n]\!]\mu$*.*

Proof Induction on n, noting that $\widehat{\mathsf{CL}}(n+1) = \widehat{\mathsf{CL}}(n)/\widehat{\mathsf{CL}}(n)$. □

3.2 The TS Hierarchy

I won't linger on the TS hierarchy, except to point out one striking thing about it: as we climb the hierarchy, a single model grows in importance for determining consequence. And at the limit TS_ω, we can forget about all other models altogether.

Fact 14 *Let* $\mathfrak{m}_{.5}$ *be the model such that* $\mathfrak{m}_{.5}(\phi) = .5$ *for every atomic sentence* ϕ*. If* $\mathfrak{m}$ *is a* TS_n *meta*^{n}*counterexample to a meta*^{n}*inference* μ*, then* $\mathfrak{m}_{.5}$ *is as well. If* $\mathfrak{m}$ *is a* TS_ω *counterexample to a minference* μ*, then* $\mathfrak{m}_{.5}$ *is as well.*

Proof By Fact 11, $\mathfrak{m}_{.5} \sqsubseteq \mathfrak{m}$. The result follows from this by Fact 12. □

Corollary 4 *For any* $m \leq n$*, a meta*^{m}*inference is* $\widehat{\mathsf{TS}_n}$ *valid iff the single model* $\mathfrak{m}_{.5}$ *is not a* $\widehat{\mathsf{TS}_n}$ *counterexample to it. For any* m*, a meta*^{m}*inference is* TS_ω *valid iff* $\mathfrak{m}_{.5}$ *is not a* TS_ω *counterexample to it.*

[30] Compare [71, Fact 1], [43, Lemma 20].

3.3 The ST Hierarchy

Here, I look in particular at the ST_ns and ST_ω, drawing out facts that apply only on this side of the hierarchy.

Fact 15 [43, Lemma 21] *For any n:* $\widehat{\mathsf{ST}_n} \approx_n \widehat{\mathsf{CL}}$

Proof We need to show that $\mathcal{C}(\mathsf{ST}_n) = \mathcal{C}(\widehat{\mathsf{CL}}(n))$.

LTR: Any $\widehat{\mathsf{CL}}(n)$ meta^{n}counterexample to a meta^{n}inference is also an ST_n meta^{n}counterexample to it, by Fact 13.

RTL: Take any ST_n meta^{n}counterexample $\mathfrak{m}$ to a meta^{n}inference. Consider any model $\mathfrak{m}'$ such that for all atomic sentences ϕ: $\mathfrak{m}'(\phi) \in \{1, 0\}$ and if $\mathfrak{m}(\phi) \in \{1, 0\}$ then $\mathfrak{m}'(\phi) = \mathfrak{m}(\phi)$. By Fact 11, $\mathfrak{m} \sqsubseteq \mathfrak{m}'$, and so by Fact 12, $\mathfrak{m}'$ is a ST_n meta^{n}counterexample to the same meta^{n}inference. Since for all atomic ϕ, we have $\mathfrak{m}'(\phi) \in \{1, 0\}$, we can show by induction on ψ that for all sentences ψ we have $\mathfrak{m}'(\psi) \in \{1, 0\}$; that is, $\mathfrak{m}'$ is 2-valued. Since $\mathfrak{m}'$ is 2-valued and an ST_n meta^{n}counterexample, it is also a $\widehat{\mathsf{CL}}(n)$ meta^{n}counterexample, by Fact 13. □

Corollary 5 *For any m, n:* $\widehat{\mathsf{ST}_n} \approx_{\min(m,n)} \widehat{\mathsf{ST}_m}$

Proof Wlog, let $m \leq n$. By Fact 15 we have $\widehat{\mathsf{ST}_m} \approx_m \widehat{\mathsf{CL}}$ and $\widehat{\mathsf{ST}_n} \approx_n \widehat{\mathsf{CL}}$. Since $\widehat{\mathsf{ST}_n}$ and $\widehat{\mathsf{CL}}$ are downward coherent by Fact 2, then by Fact 7 we have $\widehat{\mathsf{ST}_n} \approx_m \widehat{\mathsf{CL}}$. And since $\approx_m$ is symmetric and transitive, this gives $\widehat{\mathsf{ST}_m} \approx_m \widehat{\mathsf{ST}_n}$. □

Corollary 6 [43, Thm. 23] $\mathsf{ST}_\omega \approx \widehat{\mathsf{CL}}$

Proof Immediate from Fact 15. □

So we have $\mathcal{C}(\mathsf{ST}_\omega) = \mathcal{C}(\widehat{\mathsf{CL}})$; ST_ω gives us exactly the same full consequence relation as $\widehat{\mathsf{CL}}$ does. Despite this matching, the underlying models remain very different. ST_ω uses the full space of strong Kleene models, while $\widehat{\mathsf{CL}}$ ignores all but the two-valued models. As a result, ST_ω inherits the flexibility provided by strong Kleene models. For example, just as we saw for ST_0 in Section 1.5.2, ST_ω too can be nontrivially extended to a first-order language with a transparent truth predicate, and now while maintaining the validity of every $\widehat{\mathsf{CL}}$ valid minference at all levels.

4 An Objection to ST

The existence of the ST_n hierarchy, and perhaps especially of ST_ω, has been seen to pose a challenge to advocates of ST. Since ST is just ST_0, why stop there? In particular, if we're seeking agreement with classical logic, why wouldn't we want to pursue the additional agreement with $\widehat{\mathsf{CL}}$ achieved by higher levels of this hierarchy, or even the total agreement achieved by ST_ω?

Something like this objection to ST has been pressed by those who have explored this hierarchy. Here's [34, emphasis both removed and added, notation changed]:[31]

> Non-classical theories of truth pursue two conflicting desiderata. On the one hand, they search for a paradox-free transparent truth predicate. On the other hand, *they want to retain as much classical logic as possible*.... Thus, though it might be argued that ST seems to do much better than the other inferential non-classical solutions to paradoxes—precisely because it resolves paradoxes while 'mutilating' less classical logic than the other non-classical theories, ST_1 *seems to work even better than* ST. ST_1 retains every classically valid inference, as ST does, but, moreover, it recovers every classically valid metainference—while ST loses Cut (and many other classically valid metainferences) (19).

And here is [43, notation changed]:

> [T]he proponent of logics like [ST] as solutions to the paradoxes faces some difficult questions. First, they must say whether or not they mean to generalize their view to higher finite levels. If they don't, they must explain why the 'more classical logic is better' line of thought... is misguided (368).[32]

As someone who has defended ST and some of its relatives as providing valuable approaches to paradoxes,[33] I was initially pulled by this kind of objection. But I've come to think that that was a mistake, and in this section I explain why.

4.1 What's Being Objected To

First, it's important to clarify the target of the objection. The counterexample relations ST and $\widehat{\mathsf{ST}}$ are distinct: the first says only when a model is a counterexample to a meta^{0}inference, while the second says when a model is a counterexample to a meta$^\ell$inference for any level ℓ. As far as I know, nobody has so far put forward any endorsement of $\widehat{\mathsf{ST}}$, only of ST. And as I've pointed out above, an advocate of ST as a useful meta^{0}counterexample relation has thereby taken on no commitments at all regarding meta^{n}counterexample relations for $n \geq 1$.

The quoted passages, however, seem to assume that any advocate of ST in fact means to advocate $\widehat{\mathsf{ST}}$ as well. They then try to push for adopting $\widehat{\mathsf{ST}_1}$, $\widehat{\mathsf{ST}_2}$, ..., or perhaps implicitly ST_ω instead of $\widehat{\mathsf{ST}}$, on the grounds that these full counterexample relations agree with $\widehat{\mathsf{CL}}$ at higher and higher levels. But the assumption is unwarranted, and this means that the objection fails immediately as an objection to ST. After all, $\mathsf{ST} = \widehat{\mathsf{ST}_1}(0) = \widehat{\mathsf{ST}_2}(0) = \ldots = \mathsf{ST}_\omega(0)$. So one thing all these full counterexample relations *agree* on is that ST gives the correct story for meta^{0}counterexamples. So if, say, ST_ω is correct, then the advocate of ST turns out

[31] See also the concluding section of [35].

[32] This passage as written talks about $\widehat{\mathsf{ST}}$, not ST. I've substituted ST because I think this is a typo, since it's not clear what it would mean to 'generalize $\widehat{\mathsf{ST}}$ to higher levels'. After all, $\widehat{\mathsf{ST}}$ is already a full consequence relation. ([43] uses '$\mathbf{T}_1$' for $\widehat{\mathsf{ST}}$ and 'T_1' for ST, so it would be a small typo.)

[33] in for example [65, 68]

to be *right*: ST_ω goes beyond ST, but they match perfectly everywhere ST has anything to say. The disagreements between the full counterexample relations under consideration are all at higher levels, where ST makes no pronouncements.

The real objection here, then, is not to ST at all, but to $\widehat{\mathsf{ST}}$—and again, as far as I know, nobody has yet endorsed $\widehat{\mathsf{ST}}$ for any purpose. However, the objection is still worth thinking through, to see if it perhaps gives a good reason not to endorse $\widehat{\mathsf{ST}}$. In the rest of this paper, I argue that it does not; any hypothetical or future advocates of $\widehat{\mathsf{ST}}$ should not be swayed by this objection.

4.2 Why Be Classical?

One initial reason to be suspicious of the objection is that it relies on the claim that 'more classical logic is better'. But nobody should seek to be classical just for classicality's sake. Classical logic has had detractors for as long as it has existed, and many of those detractors have had very good reasons for their worries. It's true that classical logic gained a certain sort of hegemonic status in some philosophical communities in the late 20th century, but that moment (fortunately) seems to be past us now.[34]

Classical logic is an inheritance we've received, not a goal we're aiming for. Like any cultural inheritance, it contains many different strands: its two-valued model theory, its Boolean model theory, various proof systems, philosophical commitments to bivalence or noncontradiction (which themselves take various forms), multiple distinct second- and higher-order formulations, and so on. We might well want to hold to some of these while abandoning others. The objection in question, though, depends on a very particular understanding of 'classical logic': $\widehat{\mathsf{CL}}$. It also depends on a very particular understanding of what 'more classical logic' is: full agreement with $\mathcal{C}(\widehat{\mathsf{CL}}(\ell))$ for higher and higher levels ℓ. We need to ask whether the reasons often given in support of 'classical logic' fit with these very particular understandings, and—so understood—whether they fit with the idea that 'more classical logic is better'. In fact, I know of no defense of classical logic that does.

Most applications of classical logic, and so most defenses of it, are based on one of two aspects of it: its two-valued model theory,[35] or its meta^{0}consequence relation. For an example of the first kind of defense, consider [60, p. 186, emphasis in original]:

> [C]lassical semantics and logic are vastly superior to the alternatives in simplicity, power, past success, and integration with other theories in other domains. It would not be wholly unreasonable to insist on these grounds alone that bivalence *must* somehow apply to vague utterances…

('Classical semantics' here means specifically familiar 2-valued model theory, on my reading of this passage.) And for an example of the second kind of defense, consider [50, p. 223]:

[34] For useful remarks written while that moment was ongoing, see for example [42, Introduction], [36]; [37].

[35] Although I focus on two-valued models here, work on Boolean-valued models as in [5, Ch. 1] plays out the same for present purposes; note that the three-valued models used throughout this paper are not Boolean.

> [M]y definition yields a classical consequence relation, and this is important... An important constraint on a definition of validity is that it counts intuitively valid forms of reasoning as valid—and the classically valid inference forms are all prima facie paradigms of valid reasoning...

Arguments for classical logic that proceed in either of these ways, though, cannot fit with Pailos's and Scambler's objections. This is because these objections focus on similarities between $\widehat{\mathsf{ST}_1}$ or ST_ω and $\widehat{\mathsf{CL}}$ that are not shared with $\widehat{\mathsf{ST}}$. But the use of two-valued models distinguishes $\widehat{\mathsf{CL}}$ alone out of these: all of $\widehat{\mathsf{ST}}$, $\widehat{\mathsf{ST}_1}$, and ST_ω make use of the full range of strong Kleene models. So no argument for $\widehat{\mathsf{CL}}$ based on two-valued models could fit the bill for Pailos and Scambler, since $\widehat{\mathsf{ST}_1}$ and ST_ω are not any more similar to $\widehat{\mathsf{CL}}$ on this count than $\widehat{\mathsf{ST}}$ is. All three are equally dissimilar from $\widehat{\mathsf{CL}}$ in this regard.

A focus on the meta^{0}consequence relation determined also cannot distinguish any of these, as they all yield the exact same meta^{0}consequence relation. So no argument for $\widehat{\mathsf{CL}}$ based on its meta^{0}consequence relation could fit the bill for Pailos and Scambler either, since $\widehat{\mathsf{ST}_1}$ and ST_ω are again not any more similar to $\widehat{\mathsf{CL}}$ on this count than $\widehat{\mathsf{ST}}$ is. The consequence relations $\mathcal{C}(\widehat{\mathsf{CL}}(0))$, $\mathcal{C}(\widehat{\mathsf{ST}}(0))$, $\mathcal{C}(\widehat{\mathsf{ST}_1}(0))$, and $\mathcal{C}(\mathsf{ST}_\omega(0))$ are all identical.[36]

Existing defenses of classical logic, then, do not at all fit with the idea that 'more classical logic is better', where 'more classical logic' is understood as full agreement with $\mathcal{C}(\widehat{\mathsf{CL}}(\ell))$ for higher and higher levels ℓ. So what *would* it take to support $\widehat{\mathsf{CL}}$ in a way that fits with Pailos's and Scambler's objections? To be compatible with the objection's focus on consequence relations, such support should be based on consequence relations, and not on particular selections of models. And to tell the difference between $\widehat{\mathsf{ST}}$ and $\widehat{\mathsf{ST}_1}$ or ST_ω, such support must consider meta^{n}inferences for $n \geq 1$. For example, someone might argue that all meta^{1}inferences of the form $[[\Gamma \succ \Delta, \phi], [\phi, \Gamma \succ \Delta] \succ [\Gamma \succ \Delta]]$ ought to be validated. As all such meta^{1}inferences are valid in $\mathcal{C}(\widehat{\mathsf{CL}})$, in $\mathcal{C}(\widehat{\mathsf{ST}_1})$, and in $\mathcal{C}(\mathsf{ST}_\omega)$, but not all in $\mathcal{C}(\widehat{\mathsf{ST}})$, a defense of $\widehat{\mathsf{CL}}$ on *these* grounds would fit with the objection.[37]

While such a defense of $\widehat{\mathsf{CL}}$ would indeed fit with the objections I'm considering here, it would do more than this: it would replace them entirely. An argument for the validity of a minference that is invalid in $\widehat{\mathsf{ST}}$ is just directly an objection to $\widehat{\mathsf{ST}}$, already, on its own. There's no need for any detour via closeness to $\widehat{\mathsf{CL}}$. The argumentative situation around any such objection to $\widehat{\mathsf{ST}}$, then, should turn purely on how well-supported the validity of that minference is or isn't. The 'more classical logic is better' objection pressed by Pailos and Scambler simply falls by the wayside.

---

[36] This focus on the resulting meta^{0}consequence relation is present in a number of defenses of ST, for example [40, p. 156]:

> [Classicality] is the core of the advantages of the ST approach over [nonclassical approaches to truth]. There is no need, from an ST-based perspective, ever to criticize (on logical grounds) any classically-valid inference. As such, there is no need for the 'classical recapture' that so exercises many non-classical theorists.

(Thanks to an anonymous referee for pointing out this quotation.)

[37] I'll return to a similar idea in a moment, in Section 4.3.

To sum up this subsection: the objection to $\widehat{\mathsf{ST}}$ under consideration only works if we have reason to seek similarity to $\mathcal{C}(\widehat{\mathsf{CL}})$. We have reason to seek similarity to $\mathcal{C}(\widehat{\mathsf{CL}})$ only insofar as we have reason to think that $\mathcal{C}(\widehat{\mathsf{CL}})$ is getting things right. But existing reasons for thinking that 'classical logic' is getting things right aren't about $\widehat{\mathsf{CL}}$ or $\mathcal{C}(\widehat{\mathsf{CL}})$ at all. Moreover, most such reasons focus either on 2-valued models or on $\mathcal{C}(\mathsf{CL})$,[38] and so either tell against $\widehat{\mathsf{ST}}$, $\widehat{\mathsf{ST}_1}$, and ST_ω all equally; or else support all of $\widehat{\mathsf{CL}}$, $\widehat{\mathsf{ST}}$, $\widehat{\mathsf{ST}_1}$, and ST_ω equally. Either way, this puts no pressure on a hypothetical advocate of $\widehat{\mathsf{ST}}$ to adopt $\widehat{\mathsf{ST}_1}$ or ST_ω instead.

4.3 Obeying Higher Minferences

Even if the objection pressed by Pailos and Scambler fails, however, it might seem that there is an easy objection to $\widehat{\mathsf{ST}}$ in the area, one that doesn't turn on similarity to $\widehat{\mathsf{CL}}$ at all, but instead simply looks directly at plausible-seeming meta$^\ell$inferences for $\ell \geq 1$. For example, consider meta^{1}inferences of the form $[[\phi \succ \psi], [\psi \succ \rho] \succ [\phi \succ \rho]]$. These meta^{1}inferences are all in $\mathcal{C}(\widehat{\mathsf{CL}})$, $\mathcal{C}(\widehat{\mathsf{ST}_1})$, and $\mathcal{C}(\mathsf{ST}_\omega)$, but not all in $\mathcal{C}(\widehat{\mathsf{ST}})$. For example, where p, q, r are distinct atomic sentences, $[[p \succ q], [q \succ r] \succ [p \succ r]]$ is not in $\mathcal{C}(\widehat{\mathsf{ST}})$. If we had good reason to prefer counterexample relations that validate such meta^{1}inferences, then, we would have good reason to prefer $\widehat{\mathsf{ST}_1}$ or ST_ω to $\widehat{\mathsf{ST}}$.

At first blush, it might seem like we obviously have such good reason. After all, the metainferences in question seems to be a particularly simple form of 'transitivity' of consequence: they seem to say that if $[\phi \succ \psi]$ is valid and $[\psi \succ \rho]$ is valid, then $[\phi \succ \rho]$ is valid. And *that* is widely considered a minimal requirement for anything like a sensible meta^{0}consequence relation.[39] So it can seem we have a direct argument against $\widehat{\mathsf{ST}}$ and for $\widehat{\mathsf{ST}_1}$ or ST_ω, based on meta^{1}inferences like these. If simple transitivity is a desideratum, and $\widehat{\mathsf{ST}_1}$ and ST_ω have it while $\widehat{\mathsf{ST}}$ lacks it, then this gives a reason to adopt one of the former two over the latter.

In this case, though, appearances are misleading. Simple transitivity is one thing, and meta^{1}inferences of the form $[[\phi \succ \psi], [\psi \succ \rho] \succ [\phi \succ \rho]]$ are quite another. The former is a property of meta^{0}consequence relations; some are simply transitive while others are not. The latter are meta^{1}inferences, syntactic objects. The 'direct argument' above depends on conflating the two, and so it simply fails. In particular, $\mathcal{C}(\widehat{\mathsf{ST}}(0))$—which is just $\mathcal{C}(\mathsf{ST})$—*is* simply transitive; after all, it is the familiar meta^{0}consequence relation of classical logic.[40]

We should close, then, by thinking about one more relationship between adjacent levels: the relationship that, where it obtains, *would* connect meta^{1}inferences like $[[\phi \succ \psi], [\psi \succ \rho] \succ [\phi \succ \rho]]$ to simple transitivity.

[38] not $\mathcal{C}(\widehat{\mathsf{CL}})$!

[39] This is called 'simple transitivity' in [58], which, even while arguing against other commonly-accepted forms of transitivity, has "I am sympathetic to the idea that... simple transitivity... should be incorporated in any genuine notion of logical consequence" (100).

[40] Indeed, $\mathcal{C}(\mathsf{ST})$ has much stronger transitivity properties as well, as can be seen by the same reasoning. For discussion of some of these properties, see [41].

Definition 24 A meta^{n}consequence relation Σ_n *obeys* a meta^{n+1}consequence relation Σ_{n+1} iff for every $[\Gamma \succ \phi] \in \Sigma_{n+1}$, either there is some $\gamma \in \Gamma$ with $\gamma \notin \Sigma_n$ or $\phi \in \Sigma_n$.

A full consequence relation Σ is *self-obeying at level n* iff $\Sigma(n)$ obeys $\Sigma(n+1)$.

A full counterexample relation X is *self-obeying at level n* iff $\mathcal{C}(\mathsf{X})$ is, and *self-disobeying at level n* otherwise.

Interlude on Definition 24 Note that this definition uses only single-conclusion minferences. This is important! The property arrived at by generalizing this to multiple conclusions is too strong for my purposes. This interlude clarifies the situation, I hope. First, here's the stronger property we might consider:

Definition 25 A meta^{n}consequence relation Σ_n *strongly obeys* a meta^{n+1}consequence relation Σ_{n+1} iff for every $[\Gamma \succ \Delta] \in \Sigma_{n+1}$, either there is some $\gamma \in \Delta$ with $\gamma \notin \Sigma_n$ or there is some $\delta \in \Delta$ with $\delta \in \Sigma_n$.

A full consequence relation Σ is *strongly self-obeying at level n* iff $\Sigma(n)$ obeys $\Sigma(n+1)$.

Scambler [43, p. 367] says that a full consequence relation that is strongly self-obeying at all levels is 'closed under its own laws'. However, this is *not* connected to 'closure' in the ordinary sense. A *closure operation C* on a partially ordered set $\langle S, \leq \rangle$ is an operation $C : S \to S$ such that for all $x, y \in S$, we have $C(x) \leq C(y)$ iff $x \leq C(y)$.[41] Given such an operation, an $x \in S$ is *closed* iff $x = C(x)$.

Now, consider the following two very small full consequence relations: $\Sigma_a = \{[\succ p, q], p\}$, and $\Sigma_b = \{[\succ p, q], q\}$. These are strongly self-obeying at every level. And consider also $\Sigma_c = \{[\succ p, q]\}$; this is not strongly self-obeying at level -1. However:

Fact 16 *If we consider full consequence relations as ordered by $\subseteq$, then any closure operation C on full consequence relations such that Σ_a and Σ_b are both closed is also such that Σ_c is closed as well.*

Proof Since Σ_a and Σ_b are closed, they are $C(\Sigma_a)$ and $C(\Sigma_b)$, respectively. And since $\Sigma_c \subseteq C(\Sigma_a)$ and $\Sigma_c \subseteq C(\Sigma_b)$, then as C is a closure operation we have $C(\Sigma_c) \subseteq C(\Sigma_a$ and $C(\Sigma_c) \subseteq C(\Sigma_b)$; that is, $C(\Sigma_c) \subseteq C(\Sigma_a) \cap C(\Sigma_b)$. That is to say, $C(\Sigma_c) \subseteq \Sigma_a \cap \Sigma_b = \Sigma_c$. And (again, since C is a closure operation) we also have $\Sigma_c \subseteq C(\Sigma_c)$. So $\Sigma_c = C(\Sigma_c)$; that is, Σ_c is closed. □

So the property of being strongly self-obeying at every level cannot be understood as the property of being closed under *any* closure operation on full consequence relations, at least if we're considering them as being ordered by $\subseteq$. By contrast, being self-obeying at every level *is* the property of being closed under the following

[41] See for example [12, Ch. 7].

operation:

$$C(\Sigma) = \bigcap\{\Sigma' | \Sigma' \text{ is self-obeying at every level and } \Sigma \subseteq \Sigma'\}$$

(This works for self-obedience but not strong self-obedience because the former but not the latter is itself closed under intersections.) The language of being ‘closed’, then, fits with self-obedience, not with strong self-obedience.[42] By way of transition back to the main thread of the text, note that the following Fact 17 would be false for strong self-obedience, as $\mathcal{C}(\widehat{\mathsf{CL}})$, for example, is self-obeying but not strongly self-obeying at level -1.

End interlude.

Fact 17 *If a full counterexample relation* X *is* n*-upward coherent, then* X *is self-obeying at level* m *for all* $m \geq n$.

Proof Suppose that for a particular meta^{m+1}inference $[\Gamma \succ \phi]$, that $\Gamma \subseteq \mathcal{C}(\mathsf{X}(m))$ and $\phi \notin \mathcal{C}(\mathsf{X}(m))$; it suffices for the claim to show that some model is a X meta^{m+1}counterexample to $[\Gamma \succ \phi]$. Since X is n-upward coherent, we have that $\uparrow\mathsf{X}(m) = \mathsf{X}(m+1)$, so to be a meta^{m+1}counterexample to $[\Gamma \succ \phi]$ is to be a meta^{m}counterexample to ϕ without being a meta^{m}counterexample to any $\gamma \in \Gamma$. But since $\Gamma \subseteq \mathcal{C}(\mathsf{X}(m))$, no model is a meta^{m}counterexample to any $\gamma \in \Gamma$; and as $\phi \notin \mathcal{C}(\mathsf{X}(m))$, some model is a meta^{m}counterexample to ϕ. □

Interestingly, even though as we've seen $\widehat{\mathsf{ST}_n}$ is not m-upward coherent for $m < n$, it's still the case that $\widehat{\mathsf{ST}_n}$ is fully self-obeying.

Fact 18 *For all* m, n, $\widehat{\mathsf{ST}_n}$ *is self-obeying at level* m.

Proof Since $\widehat{\mathsf{ST}_n}$ is n-upward coherent by Facts 3 by 17 $\widehat{\mathsf{ST}_n}$ is self-obeying at m for all $m \geq n$. For $m < n$, however, we have $\widehat{\mathsf{ST}_n} \approx_m \widehat{\mathsf{CL}}$ and $\widehat{\mathsf{ST}_n} \approx_{m+1} \widehat{\mathsf{CL}}$, by Facts 15 and 7. By Facts 3 and 17, $\widehat{\mathsf{CL}}$ is self-obeying at every level, and so self-obeying at m. But since being self-obeying at m is just a matter of which minferences are validated at levels m and $m+1$, this means that $\widehat{\mathsf{ST}_n}$ is self-obeying at m as well. □

Corollary 7 *For all* m, ST_ω *is self-obeying at* m.

But this only goes so far. In particular, it depends on the agreement with $\widehat{\mathsf{CL}}$. Recall Section 1.5, however; in the intended applications of these strong Kleene models, the aim is to explore restricted classes of models for which $\widehat{\mathsf{ST}}$ and $\widehat{\mathsf{CL}}$ no longer agree, for applications at least to vagueness and truth.

[42] Relatedly, [68, p. 849] defines ‘metainferences’ as ‘principles under which a consequence relation might (or might not) be closed’; this closure-based way of thinking is sensible enough for the single-conclusion meta^{1}inferences considered there, but does not easily extend to multiple-conclusion minferences.

Rather than developing full theories of vagueness or truth here, it's enough for present purposes to see how things play out if we choose some particular atomic sentence λ and develop new counterexample relations that in effect restrict our models to those models $\mathfrak{m}$ for which $\mathfrak{m}(\lambda) = .5$:

Definition 26 Given a counterexample relation X, let $\mathsf{X}\lambda$ be the counterexample relation such that a model $\mathfrak{m}$ is a $\mathsf{X}\lambda$ counterexample to a minference μ iff both $\mathfrak{m}$ is a X counterexample to μ and $\mathfrak{m}(\lambda) = .5$.

It is immediate that there are no $\widehat{\mathsf{CL}\lambda}$ counterexamples to any minference (since any such counterexample would need to both be 2-valued and to assign the value .5 to λ), and so $\mathcal{C}(\widehat{\mathsf{CL}\lambda})$ is the set of all minferences. So we can really just forget about $\widehat{\mathsf{CL}\lambda}$.

Let's look, though, at what happens to the ST_n hierarchy. First, $\mathcal{C}(\widehat{\mathsf{ST}_\ell\lambda}(0))$ is no longer simply transitive, for $\ell \geq 0$; nor is $\mathcal{C}(\mathsf{ST}_\omega)$. For example, for atomic sentences p, q, distinct from each other and from λ, these meta^{0}consequence relations contain $[p \succ \lambda]$ and $[\lambda \succ q]$, but not $[p \succ q]$. If simple transitivity is a desideratum, this is a problem—but it is again a problem shared equally by $\widehat{\mathsf{ST}\lambda}$, $\widehat{\mathsf{ST}_1\lambda}$, and $\mathsf{ST}_\omega\lambda$. Moreover, none of $\mathcal{C}(\widehat{\mathsf{ST}\lambda})$, $\mathcal{C}(\widehat{\mathsf{ST}_1\lambda})$, or $\mathcal{C}(\mathsf{ST}_\omega\lambda)$ is self-obedient at level -1, since all contain λ and $[\lambda \succ \bot]$, but none contains $\bot$. If self-obedience is a desideratum, this is again a problem—and again, it is a problem shared equally by $\widehat{\mathsf{ST}\lambda}$, $\widehat{\mathsf{ST}_1\lambda}$, and $\mathsf{ST}_\omega\lambda$.

There are some interesting differences, however. We have that $\mathcal{C}(\widehat{\mathsf{ST}_1\lambda})$ and $\mathcal{C}(\mathsf{ST}_\omega\lambda)$ contain $[[p \succ \lambda], [\lambda \succ q] \succ [p \succ q]]$, while $\mathcal{C}(\widehat{\mathsf{ST}_0\lambda})$ does not contain this meta^{1}inference. With considerations of self-obedience to the fore, this difference matters: it provides an example showing that $\mathcal{C}(\widehat{\mathsf{ST}_1\lambda})$ is self-disobeying at level 0.

Indeed, as [43, Thms. 29, 30] show, for all $m < n$, $\widehat{\mathsf{ST}_n\lambda}$ is self-disobeying at level m, and $\mathsf{ST}_\omega\lambda$ is self-disobeying at all levels. We also know by Facts 17 and 3 that for all $m \geq n$, $\widehat{\mathsf{ST}_n\lambda}$ is self-obeying at level m. Climbing up the ST hierarchy thus pushes self-obedience farther and farther off into higher levels, until at the limit, ST_ω pushes self-obedience out of reach entirely.

What to make of this? Again, $\widehat{\mathsf{CL}\lambda}$ is off the table, because it's the empty counterexample relation. We're supposing that the set of meta^{0}inferences common to $\mathcal{C}(\widehat{\mathsf{ST}\lambda})$, $\mathcal{C}(\widehat{\mathsf{ST}_1\lambda})$, and $\mathcal{C}(\mathsf{ST}_\omega\lambda)$ is a good story about meta^{0}inferential validity, and trying to think about what to make of the higher-level differences. Here, if self-obedience is a desideratum for counterexample relations, then $\widehat{\mathsf{ST}\lambda}$ outperforms the others. None of these are self-obedient at level -1; unlike $\widehat{\mathsf{ST}_1\lambda}$ and $\mathsf{ST}_\omega\lambda$, however, $\widehat{\mathsf{ST}}$ is self-obedient at all other levels. Any reason to treat self-obedience as a desideratum, then, is a reason to prefer $\widehat{\mathsf{ST}\lambda}$ to these other counterexample relations. Whether or not there is such a reason is something for future work to consider.

Conversely, any reason to treat validating $[[p \succ q], [q \succ r] \succ [p \succ r]]$ as a desideratum would be a reason to prefer $\widehat{\mathsf{ST}_1\lambda}$ or $\widehat{\mathsf{ST}_\omega\lambda}$ to $\widehat{\mathsf{ST}\lambda}$. But the desire for simple transitivity provides no such reason, as none of these counterexample relations determines a simply transitive consequence relation.

5 Conclusion

The study of minferences, and mixed consequence relations on them, reveals a great deal of new texture in what seemed to be relatively familiar ground. In this paper, I've tried to develop a set of relatively useful and relatively general tools for exploring this texture. I've also showed how to apply these tools to the TS_n and ST_n hierarchies and tried to answer an objection to $\widehat{\mathsf{ST}}$ due to [35, 43], as well as a related objection based on simple transitivity. Finally, I've looked at the status of self-obedience, and pointed out that that if we want a full counterexample relation that matches $\mathsf{ST}\lambda$ at level 0 and is as self-obedient as possible, $\widehat{\mathsf{ST}\lambda}$ seems to be the way to go.

Funding Research partially supported by "Logic and Substructurality", Grant FFI2017-84805-P, Government of Spain, and by "Substructural logics for bounded resources", FT190100147, Australian Research Council.

References

1. Barrio, E., Pailos, F., & Szmuc, D. (2020a). A recovery operator for nontransitive approaches. *Review of Symbolic Logic*, *13*(1), 80–104.
2. Eduardo, A. B., Federico, P., & Damián, S (2019). (Meta)inferential levels of entailment beyond the Tarskian paradigm. Synthese. To appear.
3. Barrio, E. A., Pailos, F., & Szmuc, D. (2020b). A hierarchy of classical and paraconsistent logics. *Journal of Philosophical Logic*, *49*(1), 93–120.
4. JcBeall, & van Fraassen, B. C. (2003). *Possibilities and Paradox: An introduction to modal and many-valued logic*. Oxford: Oxford University Press.
5. Bell, J. L. (2011). *Set theory: Boolean-valued models and independence proofs* Vol. 47. Oxford: Oxford University Press.
6. Blamey, S. (2002). Partial logic. In D. M. Gabbay, & F. Guenthner (Eds.) *Handbook of philosophical logic*. 2nd edn., (Vol. 5 pp. 261–353). Springer.
7. Blasio, C., Marcos, J. ., & Wansing, H. (2017). An inferentially many-valued two-dimensional notion of entailment. *Bulletin of the Section of Logic*, *46*(3/4), 233–262.
8. Brady, R. T. (1971). The consistency of the axioms of abstraction and extensionality in a three-valued logic. *Notre Dame Journal of Formal Logic*, *12*(4), 447–453.
9. Chemla, E., & Égré, P. (2019a). From many-valued consequence to many-valued connectives. Synthese. https://doi.org/10.1007/s11229-019-02344-0. To appear.
10. Chemla, E., & Égré, P. (2019b). Suszko's problem: mixed consequence and compositionality. *Review of Symbolic Logic*, *12*(4), 736–767.
11. Chemla, E., Égré, P., & Spector, B. (2017). Characterizing logical consequence in many-valued logic. *Journal of Logic and Computation*, *27*(7), 2193–2226.
12. Brian, A., & Hilary, A. (2002). *Priestley davey introduction to lattices and order*. Cambridge: Cambridge University Press.
13. Dittrich, J. (2020). Paradox, arithmetic, and nontransitive logic. PhD thesis, Ludwig Maximilians Universität München.
14. Dunn, M. J., & Hardegree, G. M. (2001). *Algebraic methods in philosophical logic*. Oxford: Oxford University Press.
15. Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.
16. Szymon, F. (2004). Formalization of a plausible inference. *Bulletin of the Section of Logic*, *33*, 41–52.
17. Girard, J. Y. (1987). *Proof theory and logical complexity* Vol. I. Napoli: Bibliopolis.
18. Gupta, A. (1982). Truth and paradox. *Journal of Philosophical Logic*, *11*(1), 1–60.
19. Haack, S. (1974). *Deviant logic*. Cambridge: Cambridge University Press.
20. Haack, S. (1978). *Philosophy of logics*. Cambridge: Cambridge University Press.
21. Halbach, V., & Horsten, L. (2006). Axiomatizing Kripke's theory of truth. *Journal of Symbolic Logic*, *71*(2), 677–712.

22. Humberstone, L. (1988). Heterogeneous logic. *Erkenntnis*, *29*(3), 395–435.
23. Humberstone, L. (2012). *The Connectives*. Cambridge Massachusetts: MIT Press.
24. Rosanna, K. (2000). *Theories of vagueness*. Cambridge: Cambridge University Press.
25. Körner, S. (1966). *Experience and theory*. New York: Humanities Press.
26. Michael, K. (1988). Kripke and the logic of truth. *Journal of Philosophical Logic*, *17*(3), 225–278.
27. Saul, K. (1975). Outline of a theory of truth. *Journal of Philosophy*, *72*(19), 690–716.
28. Hannes, L. (1999). Truth and the liar in De Morgan-valued models. *Notre Dame Journal of Formal Logic*, *40*(4), 496–514.
29. Grzegorz, M. (1990). Q-consequence operation. *Reports on Mathematical Logic*, *24*, 49–59.
30. Martin, R. L., & Woodruff, P. W. (1976). On representing 'true-in-L' in L. In A. Kasher (Ed.) *Language in Focus: Foundations, Methods and Systems* (pp. 113–117). D.Reidel.
31. Martin-Löf, P. (1984). Constructive mathematics and computer programming. *Philosophical Transactions of the Royal Society of L. A: Mathematical, Physical and Engineering Sciences*, *312*(1522), 501–518.
32. Toby, M. (2014). Fixed points for consequence relations. *Logique et Analyse*, *227*, 333–357.
33. Nicolai, C., & Rossi, L. (2017). Principles for object-linguistic consequence: from logical to irreflexive. *Journal of Philosophical Logic*, *47*(3), 549–577.
34. Pailos, F. (2019). A family of metainferential logics. *Journal of Applied Non-Classical Logics*, *29*(1), 97–120.
35. Pailos, F. (2020). A fully classical truth theory characterized by substructural means. *The Review of Symbolic Logic*, *13*(2), 249–268.
36. Plumwood, V. (1993). The politics of reason: Towards a feminist logic. *Australasian Journal of Philosophy*, *71*(4), 436–462.
37. Priest, G. (1989). Classical logic aufgehoben. In G. Priest, R. Routley, & J. Norman (Eds.) *Paraconsistent logic: essays on the inconsistent* (pp. 131–150). Munich: Philosophia Verlag.
38. Priest, G. (2008). *An introduction to non-classical logic: From If to Is*, 2nd edn. Cambridge: Cambridge University Press.
39. Priest, G. (2013). Vague inclosures. In K. Tanaka, F. Berto, E. Mares, & F. Paoli (Eds.) *Paraconsistency: logic and applications* (pp. 367–378). Dordrecht: Springer.
40. Ripley, D. (2013). Paradoxes and failures of cut. *Australasian Journal of Philosophy*, *91*(1), 139–164.
41. Ripley, D. (2017). On the 'transitivity' of consequence relations. *Journal of Logic and Computation*, *28*(2), 433–450.
42. Routley, R., Meyer, R. K., Plumwood, V., & Brady, R.T. (1982). *Relevant Logics and their Rivals 1*. California: Ridgeview, Atascadero.
43. Scambler, C. (2020). Classical logic and the strict tolerant hierarchy. *Journal of Philosophical Logic*, *49*(2), 351–370.
44. Schütte, K. (1960). Syntactical and semantical properties of simple type theory. *The Journal of Symbolic Logic*, *25*(4), 305–326.
45. Scott, D. (1975). Combinators and classes. In *Lambda-calculus and computer science theory: proceedings of the symposium held in Rome March 25–27, 1975* (pp. 1–26). Berlin: Springer.
46. Sharvit, Y. (2017). A note on (Strawson) entailment. *Semantics & Pragmatics*, *10*(1), 1–38.
47. Shoesmith, D. J., & Smiley, T. J. (1978). *Multiple-conclusion logic*. Cambridge: Cambridge University Press.
48. Skolem, T. (1960). A set theory based on a certain three-valued logic. *Mathematica Scandinavica*, *8*(1), 127–136.
49. Skolem, T. (1963). Studies on the axiom of comprehension. *Notre Dame Journal of Formal Logic*, *4*(3), 162–170.
50. Smith, N. J. (2008). *Vagueness and degrees of truth*. Oxford: Oxford University Press.
51. Strawson, P. F. (1952). Introduction to logical theory. Methuen & Co.Ltd.
52. Tait, W. W. (1966). A nonconstructive proof of Gentzen's Hauptsatz for second order predicate logic. *Bulletin of the American Mathematical Society*, *72*(6), 980–983.
53. Tappenden, J. (1993). The liar and sorites paradoxes: Toward a unified treatment. *The Journal of Philosophy*, *90*(11), 551–577.
54. Terzian, G. (2015). Norms of truth and logical revision. *Topoi*, *34*(1), 15–23.
55. Tye, M. (1994). Sorites paradoxes and the semantics of vagueness. *Philosophical Perspectives*, *8*, 189–206.
56. Visser, A. (1984). Four valued semantics and the liar. *Journal of Philosophical Logic*, *13*(2), 181–212.

57. Von Fintel, K. (1999). NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, *16*, 97–148.
58. Weir, A. (2015). A robust non-transitive logic. *Topoi*, *34*(1), 99–107.
59. Weirich, S. (2014). Pi-forall: How to use and implement a dependently-typed language. Oregon Programming Languages Summer School. github.com/sweirich/pi-forall.
60. Williamson, T. (1994). *Vagueness*. London: Routledge.
61. Wintein, S. (2016). On all strong Kleene generalizations of classical logic. *Studia Logica*, *104*(3), 503–545.
62. Wójcicki, R. (1988). *Theory of logical calculi*. New York: Springer Science & Business Media.
63. X. (2012). Information removed for anonymous review.
64. X. (2013a). Information removed for anonymous review.
65. X. (2013b). Information removed for anonymous review.
66. X. (2013c). Information removed for anonymous review.
67. X. (2013d). Information removed for anonymous review.
68. X. (2014). Information removed for anonymous review.
69. X. (2015). Information removed for anonymous review.
70. X. (2019a). Information removed for anonymous review.
71. X. (2019b). Information removed for anonymous review.
72. Elia Z. (2008). A model of tolerance. *Studia Logica*, *90*(3), 337–368.