

Comparing Fixed-Point and Revision Theories of Truth

Philip Kremer

Received: 1 February 2004 / Accepted: 10 February 2005 / Published online: 23 June 2009
© Springer Science + Business Media B.V. 2009

Abstract In response to the liar’s paradox, Kripke developed the *fixed-point* semantics for languages expressing their own truth concepts. (Martin and Woodruff independently developed this semantics, but not to the same extent as Kripke.) Kripke’s work suggests a number of related fixed-point theories of truth for such languages. Gupta and Belnap develop their *revision theory* of truth in contrast to the fixed-point theories. The current paper considers three natural ways to compare the various resulting theories of truth, and establishes the resulting relationships among these theories. The point is to get a sense of the lay of the land amid a variety of options. Our results will also provide technical fodder for the methodological remarks of the companion paper to this one.

Keywords Fixed-point theory · Revision theory · Truth

1 Introduction

Given a first-order language \mathcal{L} , a *classical model* for \mathcal{L} is an ordered pair $M = \langle D, I \rangle$, where D , the *domain of discourse*, is a nonempty set; and where I is a function assigning to each name of \mathcal{L} a member of D , to each n -place function symbol of \mathcal{L} an n -place function on D , and to each n -place relation

P. Kremer (✉)
Department of Philosophy, University of Toronto, Toronto, Canada
e-mail: philip.kremer@utoronto.ca

symbol a function from D^n to $\{\mathbf{t}, \mathbf{f}\}$.¹ Suppose that \mathcal{L} and \mathcal{L}^+ are first-order languages, where \mathcal{L}^+ is \mathcal{L} expanded with a distinguished predicate (1-place relation symbol) \mathbf{T} , and where \mathcal{L} has a quote name ‘ A ’ for each sentence A of \mathcal{L}^+ . Then we say that \mathcal{L} and \mathcal{L}^+ are a corresponding *ground language* and *truth language*. A *ground model* for \mathcal{L} is a classical model $M = \langle D, I \rangle$ for \mathcal{L} such that $I(\text{‘}A\text{’}) = A \in D$ for each sentence A of \mathcal{L}^+ .²

Given a ground model M for \mathcal{L} , we can think of $I(X)$ as the *interpretation* or, to borrow an expression from Gupta and Belnap [3], the *signification* of X , where X is a name, function symbol or relation symbol. Gupta and Belnap characterize an expression’s or concept’s *signification* in a world w as “an abstract something that carries all the information about the expression’s [or concept’s] extensional relations in w .” If we want to interpret $\mathbf{T}x$ as “ x is true”, then, given a ground model M , we would like to find an appropriate signification, or an appropriate range of significations, for \mathbf{T} .

We might try to expand a classical ground model $M = \langle D, I \rangle$ for \mathcal{L} to a classical model $M' = \langle D, I' \rangle$ for \mathcal{L}^+ . For \mathbf{T} to express *truth*, M' should assign the same truth value to the sentences $\mathbf{T}'A$ and A , for every sentence A of \mathcal{L}^+ . Unfortunately, not every ground model $M = \langle D, I \rangle$ can thus be expanded: if λ is a (nonquote) name of \mathcal{L} and if $I(\lambda) = \neg\mathbf{T}\lambda$, then $I'(\lambda) = \neg\mathbf{T}\lambda$, so that $\mathbf{T}'\neg\mathbf{T}\lambda$ and $\mathbf{T}\lambda$ are assigned the same truth value by M' ; so $\mathbf{T}'\neg\mathbf{T}\lambda$ and $\neg\mathbf{T}\lambda$ are assigned different truth values by M' . This is a formalization of the liar’s paradox, with the sentence $\neg\mathbf{T}\lambda$ as the offending liar’s sentence.

In a semantics for languages capable of expression their own truth concepts, \mathbf{T} will not, in general, have a classical signification. Kripke [8] and Martin and Woodruff [10] present the *fixed-point* semantics for such languages. Kripke suggests a whole host of related approaches to the problem of assigning, given a ground model M , a signification to \mathbf{T} . Gupta and Belnap [3] present their *revision theories* in contrast to the various fixed-point options presented by Kripke.

In the current paper, we motivate three different ways of comparing fixed-point and revision theories of truth, and we establish the various relationships the theories have to one another in these three different senses. The general point of this is to help us get the lay of the land amid the variety of choices.

¹We will always assume that our first-order languages have countably many variables, the equals sign, connectives \neg and $\&$, and the universal quantifier \forall . The other truth-functional connectives and the existential quantifier are defined in the standard way. Our languages can have any cardinality of names, function symbols and relation symbols. We take each *predicate* to be a 1-place relation symbol.

²In various examples, below, we specify only the function symbols, relation symbols and *nonquote* names of \mathcal{L} : this fully determines both \mathcal{L} and \mathcal{L}^+ .

There is a more specific use we make of the comparisons: in the companion paper to this one, [7], we use the current results to critique on of Gupta and Belnap's motivations for their revision-theoretic approach, i.e., their claim that the revision theory has the advantage of treating truth like a classical concept when there is no vicious reference.

In the course of our investigation, we close two problems left open by Gupta and Belnap [3]. We also give a simplified proof of their "Main Lemma".

2 Fixed-Point Semantics

The intuition behind the fixed-point semantics is that pathological sentences such as the liar sentence are neither true nor false.³ In general, a *three-valued model* for a first-order language \mathcal{L} is just like a classical model, except that the function I assigns, to each n -place relation symbol, a function from D^n to $\{\mathbf{t}, \mathbf{f}, \mathbf{n}\}$. A classical model is a special case of a three-valued model. Officially \mathbf{t} (true), \mathbf{f} (false) and \mathbf{n} (either) are three truth values, but \mathbf{n} can be thought of as the absence of a truth value.⁴ We order the truth values as follows: $\mathbf{n} \leq \mathbf{n} \leq \mathbf{t} \leq \mathbf{t}$ and $\mathbf{n} \leq \mathbf{n} \leq \mathbf{f} \leq \mathbf{f}$. We say that $M = \langle D, I \rangle \leq M' = \langle D, I' \rangle$ iff $I(X) = I'(X)$ for each name or function symbol X , and $I(R)(d_1, \dots, d_n) \leq I'(R)(d_1, \dots, d_n)$ for each n -place relation symbol R and each $d_1, \dots, d_n \in D$.

Given a three-valued model $M = \langle D, I \rangle$ and an assignment s of values to the variables, the value $Val_{M,s}(t) \in D$ of each term t is defined in the standard way. The atomic formula $Rt_1 \dots t_n$ is assigned the truth value $I(R)(Val_{M,s}(t_1), \dots, Val_{M,s}(t_n))$. To evaluate composite formulas, we must have some *evaluation scheme*: for example, if A is \mathbf{f} (false) and B is \mathbf{n} (either), then we must decide whether $(A \ \& \ B)$ is \mathbf{f} or \mathbf{n} .

For classical models, we will just use the standard classical evaluation scheme, τ : If M is a classical model for \mathcal{L} and A is a sentence of \mathcal{L} , then $Val_{M,\tau}(A)$ is simply the standard truth value of A in M . For nonclassical three-valued models, we will consider the *weak Kleene scheme*, μ , and the *strong Kleene scheme*, κ . The Kleene schemes treat negation identically: $\neg \mathbf{t} = \mathbf{f}$,

³We follow the presentation in [3] of the fixed-point semantics and of the revision theory of truth. Much of this material is culled from [3] and elsewhere. Among the numbered definitions, theorems, and lemmas, those not explicitly attributed to a source are original to the current paper.

⁴We do not consider four-valued models, with the additional truth value \mathbf{b} (oth). See Visser [13, 14] and Woodruff [15].

$\neg\mathbf{f} = \mathbf{t}$, and $\neg\mathbf{n} = \mathbf{n}$. They differ in their treatment of conjunction as in the following truth table:

A	B	$A \ \& \ B$, with μ	$A \ \& \ B$, with κ
\mathbf{t}	\mathbf{t}	\mathbf{t}	\mathbf{t}
\mathbf{t}	\mathbf{f}	\mathbf{f}	\mathbf{f}
\mathbf{t}	\mathbf{n}	\mathbf{n}	\mathbf{n}
\mathbf{f}	\mathbf{t}	\mathbf{f}	\mathbf{f}
\mathbf{f}	\mathbf{f}	\mathbf{f}	\mathbf{f}
\mathbf{f}	\mathbf{n}	\mathbf{n}	\mathbf{f}
\mathbf{n}	\mathbf{t}	\mathbf{n}	\mathbf{n}
\mathbf{n}	\mathbf{f}	\mathbf{n}	\mathbf{f}
\mathbf{n}	\mathbf{n}	\mathbf{n}	\mathbf{n}

If we treat universal quantification analogously to conjunction, then for each sentence A and for the weak and strong Kleene schemes, μ and κ , we can define $Val_{M,\mu}(A)$ and $Val_{M,\kappa}(A)$: the truth value of A in M according to μ and the truth value of A in M according to κ .

We also consider van Fraassen’s *supervaluation* scheme, σ :

$$Val_{M,\sigma}(A) = \begin{cases} \mathbf{t}, \text{ if } Val_{M',\tau}(A) = \mathbf{t} \text{ for every classical } M' \geq M \\ \mathbf{f}, \text{ if } Val_{M',\tau}(A) = \mathbf{f} \text{ for every classical } M' \geq M \\ \mathbf{n}, \text{ otherwise.} \end{cases}$$

Note: If $Val_{M,\mu}(A) = Val_{M,\kappa}(A) = Val_{M,\sigma}(A) = \mathbf{n}$, then $Val_{M,\mu}(A \vee \neg A) = Val_{M,\kappa}(A \vee \neg A) = \mathbf{n}$, but $Val_{M,\sigma}(A \vee \neg A) = \mathbf{t}$.

For the fixed-point semantics, suppose, as in Section 1, that \mathcal{L} and \mathcal{L}^+ are a corresponding ground language and truth language. And suppose that $M = \langle D, I \rangle$ is a (classical) ground model for \mathcal{L} . We want to expand M to a three-valued model by adding a signification for the predicate \mathbf{T} . Let an *hypothesis* be a function $h : D \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{n}\}$, and a *classical hypothesis*, a function $h : D \rightarrow \{\mathbf{t}, \mathbf{f}\}$. Hypotheses are potential significations of \mathbf{T} . Let $M + h$ be the model $M' = \langle D, I' \rangle$ for \mathcal{L}^+ , where I' and I agree on the constants of \mathcal{L} and where $I'(\mathbf{T}) = h$. Models of the form $M + h$ are *expanded* models. If we want \mathbf{T} to express *truth*, then we want to expand our ground model M to a model $M + h$ so that $Val_{M+h,\rho}(A) = Val_{M+h,\rho}(\mathbf{T} \cdot A)$, for every sentence A of \mathcal{L}^+ , where we are working with some evaluation scheme ρ . This is equivalent to the condition, $Val_{M+h,\rho}(A) = h(A)$. We will also insist that if $d \in D$ is not a sentence of \mathcal{L}^+ , then $I'(\mathbf{T})(d) = h(d) = \mathbf{f}$.

For $\rho = \tau, \mu, \kappa$, or σ , define the *jump operator* ρ_M on the set of hypotheses as follows, restricting the definition to classical hypotheses where $\rho = \tau$:

$$\begin{aligned} \rho_M(h)(A) &= Val_{M+h,\rho}(A), \text{ if } A \in S = \text{ the set of sentences of } \mathcal{L}^+ \\ \rho_M(h)(d) &= \mathbf{f}, \text{ if } d \in D - S. \end{aligned}$$

The hypotheses meeting our conditions, above, under which T expresses truth, are the *fixed points* of the jump operator ρ_M : the hypotheses h such that $\rho_M(h) = h$. The fixed points deliver, for the language \mathcal{L}^+ , models $M + h$ satisfying what M. Kremer [6] calls “the fixed point conception of truth”, according to which, as Kripke [8] put it, “we are entitled to assert (or deny) of a sentence that it is true precisely under the circumstances when we can assert (or deny) the sentence itself.”

Kripke [8] proves that each of μ_M , κ_M and σ_M has a fixed point for every ground model M . In fact, Kripke’s results are stronger. Say that $h \leq h'$ iff $h(d) \leq h'(d)$ for every $d \in D$. And say that a function \mathcal{F} on hypotheses is *monotone* iff, for all hypotheses h and h' , if $h \leq h'$ then $\mathcal{F}(h) \leq \mathcal{F}(h')$. The jump operators μ_M , κ_M and σ_M are monotone, for every ground model M . Kripke shows that each (total) monotone function \mathcal{F} on hypotheses not only has a fixed point, but a *least* fixed point, $lfp(\mathcal{F})$. Say that h and h' are *compatible* iff $h \leq h''$ and $h' \leq h''$ for some hypothesis h'' ; and that h is \mathcal{F} -*intrinsic* iff h is compatible with every fixed point of \mathcal{F} . For example, $lfp(\mathcal{F})$ is \mathcal{F} -intrinsic. Each (total) monotone function \mathcal{F} not only has a *least* fixed point, but a *greatest* intrinsic fixed point, $gfp(\mathcal{F})$, which is not in general identical to $lfp(\mathcal{F})$.

These results yield a number of plausible significations of T : the fixed points generated by your favourite evaluation scheme. Many have considered the proposal that the least fixed point yields the correct signification of T .⁵ M. Kremer [6] decisively argues that Kripke [8] does not endorse this proposal, and that this proposal misinterprets the fixed-point semantics: the fixed-point conception of truth favours no particular fixed point. Kremer emphasizes a tension between the fixed-point conception of truth and another intuition, the “supervenience of semantics”: the intuition that the interpretation of T should be determined by the interpretation of the nonsemantic names, function symbols and relation symbols.

The disagreement between a supervenience fixed-point theorist—for specificity, say a least-fixed-point theorist using the strong Kleene scheme—and a nonsupervenience fixed-point theorist (using the same evaluation scheme) can be brought out as follows. Suppose that, other than their use of T , the discourse of two communities X and Y is represented by the same ground language \mathcal{L} interpreted by the same ground model M , but that X ’s use of T is represented by the *least* fixed point, h_X and Y ’s use of T is represented by some fixed point $h_Y \neq h_X$. Let $\mathcal{L}_X = \langle \mathcal{L}^+, M + h_X, \kappa \rangle$ and $\mathcal{L}_Y = \langle \mathcal{L}^+, M + h_Y, \kappa \rangle$ be the interpreted languages spoken by X and Y . According to the least-fixed-point theorist, community X uses T to express truth in \mathcal{L}_X , but community Y does not use T to express truth in \mathcal{L}_Y —despite the fact that, in \mathcal{L}_Y , the sentences A and T^*A have the same truth value for each sentence A of \mathcal{L}^+ . According to the nonsupervenience theorist, on the other hand, communities

⁵See [1, 2, 4, 5, 9, 11, 12].

X and Y use T to express *truth* in \mathcal{L}_X and \mathcal{L}_Y , respectively, because h_X and h_Y are fixed points: each community’s use of T satisfies the necessary and, for the nonsupervenientist, sufficient conditions for T to express truth in the community’s language.

We have on board two proposals for interpreting the fixed-point semantics. On the *supervenience* proposal, the language spoken by a community is determined by its use of nonsemantic vocabulary—represented by a ground model—and the interpretation of T as *truth* is given by some particular fixed point, usually assumed to be the least fixed point. The greatest intrinsic fixed point might also seem natural: “The largest intrinsic fixed point is the unique ‘largest’ interpretation of Tx which is consistent with our intuitive idea of truth and makes no arbitrary choices in truth assignments. It is thus an object of special theoretical interest.” (Kripke [8].) On the *nonsupervenience* proposal, the language spoken by the community is not determined by its use of nonsemantic vocabulary: the communities X and Y , in the preceding paragraph, speak distinct languages in which T expresses truth, despite a shared ground model. If we fix an evaluation scheme and a ground model, any fixed point provides an acceptable signification of *truth*.

We will not adjudicate between these two proposals. Rather, we will introduce a number of *supervenience* theories of truth, which depend on which evaluation scheme we use, and on whether we privilege the least fixed point or the greatest intrinsic fixed point. One reason to restrict ourselves to the supervenience approach is that Gupta and Belnap’s revision theories depend on the supervenience of semantics, and so it is the supervenience fixed-point theories that are most readily comparable to the revision theories.

Definition 2.1 Let $\rho = \mu, \kappa$ or σ . The sentence A of \mathcal{L}^+ is *valid in the ground model M according to (the theory) $\mathbf{T}^{lfp,\rho}$* iff $lfp(\rho_M)(A) = \mathbf{t}$. The sentence A of \mathcal{L}^+ is *valid in the ground model M according to (the theory) $\mathbf{T}^{gifp,\rho}$* iff $gifp(\rho_M)(A) = \mathbf{t}$. We define the set of sentences *valid in M* according to such and such a theory as follows:

$$\begin{aligned} \mathbf{V}_M^{lfp,\rho} &=_{\text{df}} \{A : lfp(A) = \mathbf{t}\} \\ &= \{A : A \text{ is valid in } M \text{ according to } \mathbf{T}^{lfp,\rho}\}, \text{ and} \\ \mathbf{V}_M^{gifp,\rho} &=_{\text{df}} \{A : gifp(A) = \mathbf{t}\} \\ &= \{A : A \text{ is valid in } M \text{ according to } \mathbf{T}^{gifp,\rho}\}. \end{aligned}$$

Before we consider revision theories, we introduced two variants, defined by Kripke [8], of the supervaluation jump operator, σ_M . Say that a hypothesis h is *weakly consistent* iff the set $\{A \in S : h(A) = \mathbf{t}\}$ is consistent. (Here and

elsewhere we follow [3] and use S for the set of all sentences of \mathcal{L}^+ .) Say that h is *strongly consistent* iff $\{A \in S : h(A) = \mathbf{t}\} \cup \{\neg A : A \in S \ \& \ h(A) = \mathbf{f}\}$ is consistent. Note: a classical hypothesis is strongly consistent iff $\{A \in S : h(A) = \mathbf{t}\}$ is both consistent and complete. The jump operators $\sigma 1$ and $\sigma 2$ are defined for weakly and strongly consistent hypotheses, respectively, as follows:

$$\sigma 1_M(h)(A) = \begin{cases} \mathbf{t}, & \text{if } \tau_M(h')(A) = \mathbf{t} \text{ for every weakly consistent classical } h' \geq h \\ \mathbf{f}, & \text{if } \tau_M(h')(A) = \mathbf{f} \text{ for every weakly consistent classical } h' \geq h \\ \mathbf{n}, & \text{otherwise, for sentences } A \in S. \end{cases}$$

$$\sigma 1_M(h)(d) = \mathbf{n}, \text{ for } d \in D - S.$$

$$\sigma 2_M(h)(A) = \begin{cases} \mathbf{t}, & \text{if } \tau_M(h')(A) = \mathbf{t} \text{ for every strongly consistent classical } h' \geq h \\ \mathbf{f}, & \text{if } \tau_M(h')(A) = \mathbf{f} \text{ for every strongly consistent classical } h' \geq h \\ \mathbf{n}, & \text{otherwise, for sentences } A \in S. \end{cases}$$

$$\sigma 2_M(h)(d) = \mathbf{n}, \text{ for } d \in D - S.^6$$

The operator $\sigma 1_M [\sigma 2_M]$ is a monotone operator on the weakly [strongly] consistent hypotheses. This suffices for $\sigma 1_M [\sigma 2_M]$ to have both a least fixed point and a greatest intrinsic fixed point. We will treat $\sigma 1$ and $\sigma 2$ as two new three-valued evaluation schemes. Theories $\mathbf{T}^{lfp,\sigma 1}$, $\mathbf{T}^{gifp,\sigma 1}$, $\mathbf{T}^{lfp,\sigma 2}$, and $\mathbf{T}^{gifp,\sigma 2}$, and sets $\mathbf{V}_M^{lfp,\sigma 1}$, $\mathbf{V}_M^{gifp,\sigma 1}$, $\mathbf{V}_M^{lfp,\sigma 2}$, and $\mathbf{V}_M^{gifp,\sigma 2}$ are introduced as in Definition 2.1, above.

Kripke [8] uses the least fixed point and the greatest intrinsic fixed point to define certain properties of sentences. Fix an evaluation scheme ρ , a ground model $M = \langle D, I \rangle$ for \mathcal{L} , and a sentence A of \mathcal{L}^+ . We say that A is ρ -grounded in M iff $lfp(\rho_M)(A) \neq \mathbf{n}$, and ρ -intrinsic in M iff $gifp(\rho_M)(A) \neq \mathbf{n}$. Suppose that a, b and c are names in \mathcal{L} , and that $I(a) = \neg \mathbf{T}a$, $I(b) = \mathbf{T}b$, and $I(c) = (\mathbf{T}c \vee \neg \mathbf{T}c)$. Thus $\neg \mathbf{T}a$ is a liar and $\mathbf{T}b$ is a truth-teller. The liar is neither ρ -grounded nor ρ -intrinsic since it gets the value \mathbf{n} at every fixed point h . The truth-teller is neither ρ -grounded nor ρ -intrinsic since it gets the value \mathbf{t} at some fixed points, the value \mathbf{f} at others, and the value \mathbf{n} at others. $(\mathbf{T}c \vee \neg \mathbf{T}c)$ is, for example, μ - and κ -intrinsic and σ -grounded, but neither μ - nor κ -grounded: $gifp(\mu_M)(\mathbf{T}c \vee \neg \mathbf{T}c) = gifp(\kappa_M)(\mathbf{T}c \vee \neg \mathbf{T}c) = lfp(\sigma_M)(\mathbf{T}c \vee \neg \mathbf{T}c) = \mathbf{t}$, while $lfp(\mu_M)(\mathbf{T}c \vee \neg \mathbf{T}c) = lfp(\kappa_M)(\mathbf{T}c \vee \neg \mathbf{T}c) = \mathbf{n}$.

⁶An equivalent definition of $\sigma 2_M(h)(A)$ is this: $\sigma 2_M(h)(A) = \mathbf{t} [\mathbf{f}]$ iff A is true [false] in all classical models $M' \geq M + h$ such that the extension of \mathbf{T} in M' is complete and consistent. Gupta and Belnap [3] define a jump operator σ_M^c in this way, but for weakly rather than strongly consistent h . Unfortunately, the weak consistency of h does not guarantee the existence of a model $M' \geq M + h$ such that the extension of \mathbf{T} in M' is complete and consistent. In fact, the existence of such a model M' is equivalent to the strong consistency of h . The jump operator $\sigma 2_M$ is identical to σ_M^c , with the definition in [3] corrected so that it is restricted to strongly consistent h .

3 Revision Theories of Truth

Gupta and Belnap's most interesting objection to the fixed-point semantics stems from an uncommon take on a common observation: the observation that there are connectives that fixed-point languages cannot express, for example, *exclusion negation*, for which $\neg \mathbf{n} = \mathbf{t}$; and the Łukasiewicz biconditional, for which $(\mathbf{n} \equiv \mathbf{n}) = \mathbf{t}$.⁷ Their objection is *not* that there is a gap between the resources of the object language and the metalanguage, but that "there is a gap between the resources of the language that is the original object of investigation and those of the languages that are amenable to fixed point theories." (p. 101) The language that is the original object of investigation can express genuinely paradoxical sentences, whose behaviour is unstable. And one *source* of the language's ability to express such paradoxicalities is the fact that it *can* express, for example, exclusion negation. A fixed point language cannot, in the end, express genuinely paradoxical sentences: even the liar behaves stably. So fixed-point theories do not deliver an analysis of the unstable phenomenon that we are trying to understand: "There are appearances of the Liar here, but they deceive." (p. 96)

Working with a purely two-valued object language, Gupta and Belnap imagine beginning with a classical hypothesis h regarding the extension of \mathbf{T} , and then revising h by using the classical jump operator, the "rule of revision", τ_M . As the revision procedure proceeds $(h, \tau_M(h), \tau_M^2(h), \dots)$, a liar sentence will flip back and forth between \mathbf{t} and \mathbf{f} . A truth-teller will keep whatever value it had to begin with. Other sentences might display unstable behaviour to begin with, but eventually settle down to a particular value. Some sentences will be very well behaved: they will settle down to a truth value that is independent of the initial hypothesis. Gupta and Belnap formalize the carrying out of such procedures into the transfinite, with their notion of a *revision sequence*.

Given any function \mathcal{F} on hypotheses, an \mathcal{F} -sequence, or a *revision sequence* for \mathcal{F} , is an ordinal-length sequence \mathcal{S} of hypotheses such that $\mathcal{S}_{\alpha+1} = \mathcal{F}(\mathcal{S}_\alpha)$, for every ordinal α ; and every limit ordinal λ , every truth value \mathbf{x} , and every $d \in D$, we have

$$\begin{aligned} \mathcal{S}_\lambda(d) = \mathbf{x} & \text{ if there is a } \beta < \lambda \text{ such that} \\ & \mathcal{S}_\alpha(d) = \mathbf{x} \text{ for every ordinal } \alpha \text{ with } \beta \leq \alpha < \lambda. \end{aligned}$$

This second clause is the *limit* rule for \mathcal{F} -sequences. Note that if \mathcal{S} is an \mathcal{F} -sequence, then \mathcal{F} is defined on \mathcal{S}_α for every ordinal α ; so, if \mathcal{S} is a τ_M -sequence, then \mathcal{S}_α is classical for every ordinal α . Any ordinal-length sequence \mathcal{S} of hypotheses *culminates* in h iff there is an ordinal β such that $\mathcal{S}_\alpha = h$ for

⁷In the strong and weak Kleene schemes, $(\mathbf{n} \equiv \mathbf{n}) = \mathbf{n}$.

every ordinal $\alpha \geq \beta$. For the purposes of the revision theory of truth, we are primarily interested in τ_M -sequences, but other revision sequences are useful. Note that if $\rho = \mu, \kappa, \sigma, \sigma 1$, or $\sigma 2$ and if M is a ground model, then there is a unique ρ_M -sequence \mathcal{S} such that $\mathcal{S}_0(d) = \mathbf{n}$ for every $d \in D$. Furthermore, \mathcal{S} culminates in $\text{lfp}(\rho_M)$.

As mentioned, Gupta and Belnap want to formalize the behaviour of truth, instabilities and all. Relative to a ground model M , this behaviour is arguably represented by the class of τ_M -sequences: this class delivers a verdict on which sentences are well-behaved or ill-behaved, as well as a representation of *how* various sentences are ill-behaved. For this reason, Gupta and Belnap propose that the signification of truth is the revision rule τ_M , since this rule arguably fits the Gupta-Belnap characterization (see Section 1, above) of an expression's or concept's signification.

The *most* well-behaved sentences are those that are stably **t** in every τ_M -sequence: formally, a sentence A of \mathcal{L}^+ is *stably t [f]* in the τ_M -sequence \mathcal{S} iff there is an ordinal β such that for every $\gamma \geq \beta$, we have $\mathcal{S}_\gamma(A) = \mathbf{t} [\mathbf{f}]$. Accordingly, Gupta and Belnap introduce the revision theory \mathbf{T}^* :

Definition 3.1 [3] Suppose that M is a ground model for the ground language \mathcal{L} . The sentence A of \mathcal{L}^+ is *valid in M according to* (the theory) \mathbf{T}^* iff A is stably **t** in every τ_M -sequence. $\mathbf{V}_M^* =_{\text{df}} \{A \in S : A \text{ is stably true in every } \tau_M\text{-sequence}\}$.

We might want to weaken this condition on the validity of a sentence A of \mathcal{L}^+ in a ground model M . In some ground models, there are sentences that are *nearly stably t* in the following sense: they are stably **t** except possibly at limit ordinals and for a finite number of steps after limit ordinals.⁸ Formally, a sentence A of \mathcal{L}^+ is *nearly stably t [f]* in the τ_M -sequence \mathcal{S} iff there is an ordinal β such that for every $\gamma \geq \beta$, there is a natural number m such that for every $n \geq m$, we have $\mathcal{S}_{\gamma+n}(A) = \mathbf{t} [\mathbf{f}]$. Gupta and Belnap's theory $\mathbf{T}^\#$ is based on near stability:

Definition 3.2 [3] Suppose that M is a ground model for the ground language \mathcal{L} . The sentence A of \mathcal{L}^+ is *valid in M according to* (the theory) $\mathbf{T}^\#$ iff A is nearly stably **t** in every τ_M -sequence. $\mathbf{V}_M^\# =_{\text{df}} \{A \in S : A \text{ is nearly stably true in every } \tau_M\text{-sequence}\}$.

Finally, we might put constraints on which hypotheses are legitimate hypotheses concerning the extension of \mathbf{T} , and hence on which τ_M -sequences are legitimate revision sequences. A natural condition to put on the legitimacy of a classical hypothesis h is that the resulting extension of \mathbf{T} be consistent and

⁸See Example 5.7, below.

complete, i.e., that h be strongly consistent. A τ_M -sequence \mathcal{S} is *maximally consistent* iff \mathcal{S}_α is strongly consistent for every ordinal α . Gupta and Belnap’s theory \mathbf{T}^c is based on maximally consistent τ_M -sequences.

Definition 3.3 [3] Suppose that M is a ground model for the ground language \mathcal{L} . The sentence A of \mathcal{L}^+ is *valid in M according to* (the theory) \mathbf{T}^c iff A is stably \mathbf{t} in every maximally consistent τ_M -sequence. $\mathbf{V}_M^c =_{df} \{A \in S : A \text{ is stably true in every maximally consistent } \tau_M\text{-sequence}\}$.

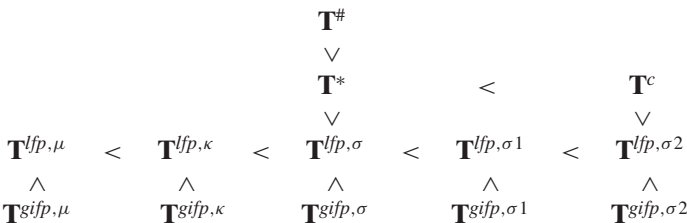
All three revision theories are *supervenience* theories in the sense of Section 2: the behaviour of truth and the status of various sentences is determined by the nonsemantic vocabulary, whose use is represented by the ground model. There is no other way to go in the revision-theoretic setting: for most ground models M , there is no class H of privileged hypotheses, like the fixed points, such that for distinct $h, h' \in H$, we could take the expanded models $M + h$ and $M + h'$ to represent distinct *languages* in which \mathbf{T} represents truth. On the revision theories, each language is represented by a ground model, and the behaviour of truth is represented by the various ways in which one hypothesis leads to another as we carry out the revision process.

4 Three Ways to Compare Theories of Truth

(The harder parts of the proofs of the theorems in this section are reserved for Section 5.) We have on the table thirteen theories of truth: ten fixed-point theories, $\mathbf{T}^{lfp,\mu}, \mathbf{T}^{lfp,\kappa}, \mathbf{T}^{lfp,\sigma}, \mathbf{T}^{lfp,\sigma 1}, \mathbf{T}^{lfp,\sigma 2}, \mathbf{T}^{gifp,\mu}, \mathbf{T}^{gifp,\kappa}, \mathbf{T}^{gifp,\sigma}, \mathbf{T}^{gifp,\sigma 1}, \mathbf{T}^{gifp,\sigma 2}$; and three revision theories, $\mathbf{T}^*, \mathbf{T}^\#, \mathbf{T}^c$. The first relation we define to compare these thirteen theories is the most obvious:

Definition 4.1 Given any two theories \mathbf{T} and \mathbf{T}' among our thirteen, we say that $\mathbf{T} \leq_1 \mathbf{T}'$ iff for every ground language \mathcal{L} , every ground model M for \mathcal{L} , and every sentence A of \mathcal{L}^+ , if A is valid in M according to \mathbf{T} then A is valid in M according to \mathbf{T}' . We say that $\mathbf{T} <_1 \mathbf{T}'$ iff $\mathbf{T} \leq_1 \mathbf{T}'$ and $\mathbf{T} \neq \mathbf{T}'$. Note that \leq_1 is reflexive and transitive.

Theorem 4.2 $<_1$ behaves as in the following diagram, i.e., it is the smallest transitive relation satisfying the conditions given in the diagram. Since \leq_1 is reflexive, the diagram completely determines \leq_1 . The subscripted 1 has been dropped from the diagram.



Proof For $\mathbf{T}^{lfp,\mu} \leq_1 \mathbf{T}^{lfp,\kappa} \leq_1 \mathbf{T}^{lfp,\sigma} \leq_1 \mathbf{T}^{lfp,\sigma_1} \leq_1 \mathbf{T}^{lfp,\sigma_2}$, it suffices to show that $lfp(\mu_M) \leq lfp(\kappa_M) \leq lfp(\sigma_M) \leq lfp(\sigma_1_M) \leq lfp(\sigma_2_M)$, for any ground model $M = \langle D, I \rangle$. For $\rho = \mu, \kappa, \sigma, \sigma_1$, or σ_2 , let $\mathcal{S}(\rho)$ be the unique ρ_M -sequence such that $\mathcal{S}(\rho)_0(d) = \mathbf{n}$ for every $d \in D$. By transfinite induction, $\mathcal{S}(\mu)_\alpha \leq \mathcal{S}(\kappa)_\alpha \leq \mathcal{S}(\sigma)_\alpha \leq \mathcal{S}(\sigma_1)_\alpha \leq \mathcal{S}(\sigma_2)_\alpha$, for every ordinal α . Our desired result follows from the fact that $\mathcal{S}(\rho)$ culminates in $lfp(\rho_M)$.

For $\mathbf{T}^{lfp,\rho} \leq_1 \mathbf{T}^{gifp,\rho}$ ($\rho = \mu, \kappa, \sigma, \sigma_1$, or σ_2), it suffices to note that $lfp(\rho_M) \leq gifp(\rho_M)$, since $lfp(\rho_M)$ is intrinsic.

$\mathbf{T}^* \leq_1 \mathbf{T}^\#$ and $\mathbf{T}^* \leq_1 \mathbf{T}^c$ can be proved directly from the definitions.

To see that $\mathbf{T}^{lfp,\sigma} \leq_1 \mathbf{T}^*$, fix a ground model $M = \langle D, I \rangle$ and let \mathcal{S} be the unique σ_M -sequence such that $\mathcal{S}_0(d) = \mathbf{n}$ for every $d \in D$. Then \mathcal{S} culminates in $lfp(\sigma_M)$. And let \mathcal{S}' be any τ_M -revision sequence. By transfinite induction, $\mathcal{S}_\alpha \leq \mathcal{S}'_\alpha$, for every ordinal α . So, if $lfp(\sigma_M)(A) = \mathbf{t}$, then A is stably \mathbf{t} in \mathcal{S}' . Since \mathcal{S}' was arbitrary, if $lfp(\sigma_M)(A) = \mathbf{t}$, then A is valid in M according to \mathbf{T}^* . So $\mathbf{T}^{lfp,\sigma} \leq_1 \mathbf{T}^*$.

Similarly, $\mathbf{T}^{lfp,\sigma_2} \leq_1 \mathbf{T}^c$.

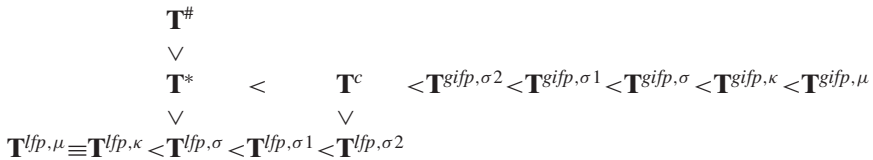
This establishes all of the positive claims of the form $\mathbf{T} \leq_1 \mathbf{T}'$ in Theorem 4.2. The counterexamples in Section 5, below, establish the negative claims of the form $\mathbf{T} \not\leq_1 \mathbf{T}'$. □

Of particular interest are ground models in which truth behaves like a classical concept. Suppose, for example, that one is devising a semantics for languages that contain their own truth predicates. All else being equal, one might want a semantics that delivers, whenever possible, something approaching a classical theory: we know that truth behaves paradoxically, but it seems an advantage to minimize this paradoxicality. Consider, for example, a classical ground model $M = \langle D, I \rangle$ that makes no distinctions, other than with quote names, among the sentences of \mathcal{L}^+ : for an extreme case, suppose that the ground language \mathcal{L} has no nonquote names, no function symbols and no nonlogical relation symbols. There is no circular reference in the ground model, and there seems to be no vicious reference of any kind. And yet $lfp(\mu_M)$ and $lfp(\kappa_M)$ are nonclassical (see the proof of Theorem 4.5): this suggests that truth does not behave like a classical concept in M , at least not according to the least-fixed-point theories $\mathbf{T}^{lfp,\mu}$ and $\mathbf{T}^{lfp,\kappa}$. On the other hand, $gifp(\mu_M)$ and $gifp(\kappa_M)$ are both classical, as is $lfp(\sigma_M)$ (this follows from Corollary 4.26, below). So, at least relative to M , the theories $\mathbf{T}^{gifp,\mu}$, $\mathbf{T}^{gifp,\kappa}$, and $\mathbf{T}^{lfp,\sigma}$ have an advantage over $\mathbf{T}^{lfp,\mu}$ and $\mathbf{T}^{lfp,\kappa}$. This motivates our definition of \leq_2 , below (Definition 4.4).

Definition 4.3 Let $\rho = \mu, \kappa, \sigma, \sigma_1$, or σ_2 . We say that $\mathbf{T}^{lfp,\rho}$ [$\mathbf{T}^{gifp,\rho}$] *dictates that truth behaves like a classical concept in the ground model M* iff $A \in \mathbf{V}_M^{lfp,\rho}$ [$\mathbf{V}_M^{gifp,\rho}$] or $\neg A \in \mathbf{V}_M^{lfp,\rho}$ [$\mathbf{V}_M^{gifp,\rho}$] for every sentence A of \mathcal{L}^+ ; equivalently, iff $lfp(\rho_M)$ [$gifp(\rho_M)$] is classical. Similarly, we say that \mathbf{T}^* [$\mathbf{T}^\#, \mathbf{T}^c$] *dictates that truth behaves like a classical concept in the ground model M* iff $A \in \mathbf{V}_M^*$ [$\mathbf{V}_M^\#, \mathbf{V}_M^c$] or $\neg A \in \mathbf{V}_M^*$ [$\mathbf{V}_M^\#, \mathbf{V}_M^c$] for every sentence A of \mathcal{L}^+ .

Definition 4.4 Given any two theories \mathbf{T} and \mathbf{T}' among our thirteen theories, we say that $\mathbf{T} \leq_2 \mathbf{T}'$ iff, for every ground language \mathcal{L} and every ground model M for \mathcal{L} , if \mathbf{T} dictates that truth behaves like a classical concept in M then so does \mathbf{T}' . Note that, among our thirteen theories, $\mathbf{T} \leq_2 \mathbf{T}'$ iff, for every ground language \mathcal{L} and every ground model M for \mathcal{L} , if \mathbf{T} dictates that truth behaves like a classical concept in M , then every sentence valid in M according to \mathbf{T} is also valid in M according to \mathbf{T}' . (This follows from Theorems 4.2 and 4.5, as noted by an anonymous referee.) We say that $\mathbf{T} \equiv_2 \mathbf{T}'$ iff both $\mathbf{T} \leq_2 \mathbf{T}'$ and $\mathbf{T}' \leq_2 \mathbf{T}$. We say that $\mathbf{T} <_2 \mathbf{T}'$ iff both $\mathbf{T} \leq_2 \mathbf{T}'$ and $\mathbf{T} \not\equiv_2 \mathbf{T}'$. Note that \leq_2 is reflexive and transitive. Note also that if $\mathbf{T} \leq_1 \mathbf{T}'$ then $\mathbf{T} \leq_2 \mathbf{T}'$.

Theorem 4.5 $<_2$ behaves as in the following diagram, i.e., it is the smallest transitive relation satisfying the conditions given in the diagram. Since \leq_2 is reflexive, the diagram completely determines \leq_2 . The subscripted 2 has been dropped from the diagram.



Proof The fact that $\mathbf{T}^{lfp,\mu} \equiv_2 \mathbf{T}^{lfp,\kappa}$ follows from the fact that neither $lfp(\mu_M)$ nor $lfp(\kappa_M)$ is classical in any ground model M . To see that $lfp(\mu_M)$ is nonclassical, let M be any ground model and let \mathcal{S} be the unique μ_M -sequence such that $\mathcal{S}_0(d) = \mathbf{n}$ for every $d \in D$. By transfinite induction, $\mathcal{S}_\alpha(\forall x(\mathbf{T}x \vee \neg \mathbf{T}x)) = \mathbf{n}$, for every ordinal α . But then $lfp(\mu_M)(\forall x(\mathbf{T}x \vee \neg \mathbf{T}x)) = \mathbf{n}$, since \mathcal{S} culminates in $lfp(\mu_M)$. Similarly, $lfp(\kappa_M)$ is nonclassical.

The following follow from the already proven part of Theorem 4.2: $\mathbf{T}^{lfp,\kappa} \leq_2 \mathbf{T}^{lfp,\sigma} \leq_2 \mathbf{T}^{lfp,\sigma^1} \leq_2 \mathbf{T}^{lfp,\sigma^2} \leq_2 \mathbf{T}^c$ and $\mathbf{T}^{lfp,\sigma} \leq_2 \mathbf{T}^* \leq_2 \mathbf{T}^\#$ and $\mathbf{T}^* \leq_2 \mathbf{T}^c$.

To see that $\mathbf{T}^c \leq_2 \mathbf{T}^{gifp,\sigma^2}$, suppose that M is a ground model in which \mathbf{T}^c dictates that truth behaves like a classical concept. So there is a classical hypothesis h in which all maximally consistent τ_M -sequences culminate. It suffices to show that h is the greatest fixed point of $\sigma 2_M$, in which case $gifp(\sigma 2_M) = h$ is classical, in which case $\mathbf{T}^{gifp,\sigma^2}$ dictates that truth behaves like a classical concept in M . Let h' be any fixed point of $\sigma 2_M$. Since h' is strongly consistent, we can choose a strongly consistent classical $h'' \geq h'$. Let \mathcal{S} be any maximally consistent τ_M -sequence with $\mathcal{S}_0 = h'' \geq h'$. By the monotonicity of $\sigma 2_M$ together with the fact that $\sigma 2_M$ agrees with τ_M on all classical hypotheses, we can show by transfinite induction that $\mathcal{S}_\alpha \geq h'$ for every ordinal α . So $h \geq h'$, since \mathcal{S} culminates in h . So h is the greatest fixed point of $\sigma 2_M$, as desired.

To see that $\mathbf{T}^{gifp,\sigma^2} \leq_2 \mathbf{T}^{gifp,\sigma^1} \leq_2 \mathbf{T}^{gifp,\sigma} \leq_2 \mathbf{T}^{gifp,\kappa} \leq_2 \mathbf{T}^{gifp,\mu}$, choose ρ and ρ' with ρ to the left of ρ' from the list $\mu, \kappa, \sigma, \sigma^1, \sigma^2$. It suffices to show that if $gifp(\rho'_M)$ is classical, then $gifp(\rho_M) = gifp(\rho'_M)$. So suppose that $gifp(\rho'_M)$ is

classical. Then it is a fixed point of τ_M and hence not only of ρ'_M but also of ρ_M . To show that $gfp(\rho_M) = gifp(\rho'_M)$, it suffices to show that $h \leq gifp(\rho'_M)$ for every fixed point h of ρ_M : this would show that $gifp(\rho'_M)$ is not only a fixed point of ρ_M but also the *greatest* fixed point of ρ_M , and hence the greatest intrinsic fixed point of ρ_M . Choose a fixed point h of ρ_M . Note that ρ'_M is defined on h ; in case ρ' is σ_1 or σ_2 , h is strongly consistent since h is a fixed point of ρ_M . Furthermore, $h = \rho_M(h) \leq \rho'_M(h)$. So there is exactly one ρ' -sequence \mathcal{S} such that $\mathcal{S}_0 = h$; moreover \mathcal{S} culminates in some fixed point h' of ρ'_M : indeed h' is the least fixed point of ρ'_M with $h \leq h'$. Since $gifp(\rho'_M)$ is classical, $gifp(\rho'_M)$ is the greatest fixed point of ρ'_M . So $h \leq h' \leq gifp(\rho'_M)$, as desired.

This establishes all of the positive claims of the form $\mathbf{T} \leq_2 \mathbf{T}'$ in Theorem 4.5. The counterexamples in Section 5, below, establish the negative claims of the form $\mathbf{T} \not\leq_2 \mathbf{T}'$. □

Remark 4.6 Theorem 4.5 answers a question of Gupta and Belnap’s ([3], Problem 6B.12): “Does the condition ‘ $lfp(\sigma_{2_M})$ is classical’ imply ‘ M is Thomason?’” (We define *Thomason* ground models below.) The answer is no, since $\mathbf{T}^{lfp,\sigma_2} \not\leq_2 \mathbf{T}^*$ (see Example 5.9, below) and since, by Theorem 4.8, below, a ground model M is Thomason iff \mathbf{T}^* dictates that truth behaves like a classical concept in M .

The next comparative relation, \leq_3 , is trickier to motivate, and is best understood in the context of investigating whether this or that theory dictates that truth behaves like a classical concept in this or that ground model M .

For starters, it is not always easy to tell whether some theory dictates that truth behaves like a classical concept in M . Gupta and Belnap devote some effort to investigating the circumstances under which, in effect, \mathbf{T}^* dictates that truth behaves like a classical concept in a ground model, though they do not put it in these terms. As we shall see, their investigation can be broadened to theories other than \mathbf{T}^* . Gupta and Belnap proceed by introducing the notion of a *Thomason* ground model, and by investigating the circumstances under which a ground model is Thomason.

Definition 4.7 [3] A ground model M is *Thomason* iff all τ_M -sequences culminate in one and the same fixed point.

Theorem 4.8 A ground model M is Thomason iff \mathbf{T}^* dictates that truth behaves like a classical concept in M .

Proof This follows immediately from the definitions. □

Gupta and Belnap’s principal results concerning Thomason models all have the same general form, and all make it relatively easy to show that a wide range of ground models are, in fact, Thomason. The simplest example concerns any ground model M for the ground language \mathcal{L} described above: a language with no nonquote names, no function symbols, and no nonlogical relation symbols.

Any such ground model is Thomason. This might be expected since, other than with quote names, there is no way to distinguish in \mathcal{L} (interpreted by M) among sentences of \mathcal{L}^+ .

This is a special case of Gupta and Belnap's result, Theorem 4.11, below. Essentially, Theorem 4.11 states that any ground model that cannot distinguish among the sentences of \mathcal{L}^+ , other than with quote names, is Thomason. First we need to make precise the notion of "distinguishing among the sentences of \mathcal{L}^+ ."

Definition 4.9 ([3], Definition 2D.2) Suppose that $M = \langle D, I \rangle$ is a ground model for the ground language \mathcal{L} and that $X \subseteq D$.

1. The interpretation of a name c is X -neutral iff $I(c) \notin X$.
2. The interpretation of an n -place relation symbol R is X -neutral iff for all $d_1, \dots, d_n, d'_i \in D$, if $d_i, d'_i \in X$ then $I(R)(d_1, \dots, d_i, \dots, d_n) = I(R)(d_1, \dots, d'_i, \dots, d_n)$.
3. The interpretation of an n -place function symbol f is X -neutral iff both
 - a. the range of $I(f)$ is disjoint from X ; and
 - b. for all $d_1, \dots, d_n, d'_i \in D$, if $d_i, d'_i \in X$ then $I(f)(d_1, \dots, d_i, \dots, d_n) = I(f)(d_1, \dots, d'_i, \dots, d_n)$.

Definition 4.10 ([3], Definition 6A.2) A ground model $M = \langle D, I \rangle$ is X -neutral iff the interpretations in M of all the nonquote names, nonlogical relation symbols, and function symbols are X -neutral.

Theorem 4.11 ([3], Theorem 6A.5) *If the ground model M is S -neutral (where S is the set of sentences of \mathcal{L}^+) then M is Thomason.*

Proof This is a special case of Corollary 4.26, below. □

Gupta and Belnap strengthen this theorem: Suppose that the ground model M can in fact distinguish among sentences of \mathcal{L}^+ , but only among sentences that are in some sense unproblematic, for example among sentences with no occurrences of \mathbf{T} or among μ -ungrounded sentences. Then M is still Thomason.

Theorem 4.12 ([3], Theorem 6B.4, Convergence to a fixed point I) *If M is X -neutral then M is Thomason provided that X contains either (i) all sentences that have occurrences of \mathbf{T} , or (ii) all sentences that are μ -ungrounded in M , or (iii) all sentences that are κ -ungrounded in M , or (iv) all sentences that are σ -ungrounded in M .*

Proof (i) is a special case of Corollary 4.26, below. (ii), (iii) and (iv) are special cases of Theorem 4.21, below. □

Note that (ii), (iii) and (iv) of Theorem 4.12 can be reworded as follows:

Theorem 4.13 *Suppose that M is a ground model, that $\mathbf{V}_M = \mathbf{V}_M^{lfp,\mu}$ or $\mathbf{V}_M^{lfp,\kappa}$ or $\mathbf{V}_M^{lfp,\sigma}$, and that $Y \subseteq \{A \in S : A \in \mathbf{V}_M \text{ or } \neg A \in \mathbf{V}_M\}$. Then if M is $(S - Y)$ -neutral, then M is Thomason.*

Gupta and Belnap present the following example as an application of Theorem 4.12. This shows how easy it can be, equipped with Theorem 4.12 or Theorem 4.13, to show that a ground model is Thomason.

Example 4.14 ([3], Example 6B.6) Suppose that the ground model $M = \langle D, I \rangle$ is S -neutral except for the name a . Furthermore suppose that Hb is true in M . Then M is Thomason if (i) $I(a) = Hb$, or (ii) $I(a) = \mathbf{T}^*Hb$, or (iii) $I(a) = Hb \vee \neg \mathbf{T}a$, or (iv) $I(a) = \mathbf{T}a \vee \neg \mathbf{T}a$.

Gupta and Belnap’s other main theorem concerning Thomason models is as follows:

Theorem 4.15 ([3], Theorem 6B.4, Convergence to a fixed point II) *Suppose that the ground model M is $(S - Y)$ -neutral and that $Y \subseteq \{A : A \in \mathbf{V}_M^* \text{ or } \neg A \in \mathbf{V}_M^*\}$. Then M is Thomason.*

Proof This is a special case of Theorem 4.21, below. □

Gupta and Belnap then go on to ask a related question:

Question 4.16 ([3], Problem 6B.15) Suppose that the ground model M is $(S - Y)$ -neutral and that $Y \subseteq \{A : A \in \mathbf{V}_M^c \text{ or } \neg A \in \mathbf{V}_M^c\}$. Is M Thomason?

As pointed out above, an investigation into the conditions under which a ground model M is Thomason is, in effect, an investigation into the conditions under which \mathbf{T}^* dictates that truth behaves like a classical concept in M . It turns out that, for a wide range of our theories \mathbf{T} , if M is $(S - Y)$ -neutral where $Y \subseteq \{A : A \in \mathbf{V}_M \text{ or } \neg A \in \mathbf{V}_M\}$ and where $\mathbf{V}_M = \{A \in S : A \text{ is valid in the ground model } M \text{ according to } \mathbf{T}\}$, then \mathbf{T}^* does, in fact, dictate that truth behaves like a classical concept in M . To help generalize this investigation, we define a third relation \leq_3 between theories.

Definition 4.17 Suppose that \mathbf{T} and \mathbf{T}' are among our thirteen theories and that, for any ground model M , $\mathbf{V}_M = \{A \in S : A \text{ is valid in the ground model } M \text{ according to } \mathbf{T}\}$. We say that $\mathbf{T} \leq_3 \mathbf{T}'$ iff for every ground language \mathcal{L} , every ground model M for \mathcal{L} , and every $Y \subseteq \{A : A \in \mathbf{V}_M \text{ or } \neg A \in \mathbf{V}_M\}$, if M is $(S - Y)$ -neutral then \mathbf{T}' dictates that truth behaves like a classical concept in M . We say that $\mathbf{T} <_3 \mathbf{T}'$ iff $\mathbf{T} \leq_3 \mathbf{T}'$ and $\mathbf{T}' \not\leq_3 \mathbf{T}$. We will see that \leq_3 is transitive but not reflexive.

Remark 4.18 Theorems 4.13 and 4.15 can be summarized as follows: $\mathbf{T}^{lfp,\mu} \leq_3 \mathbf{T}^*$, $\mathbf{T}^{lfp,\kappa} \leq_3 \mathbf{T}^*$, $\mathbf{T}^{lfp,\sigma} \leq_3 \mathbf{T}^*$, and $\mathbf{T}^* \leq_3 \mathbf{T}^*$. Question 4.16 amounts to this: $\mathbf{T}^c \leq_3 \mathbf{T}^*$? Theorem 4.21, below, delivers a negative answer to this.

Lemma 4.19 \leq_3 is transitive.

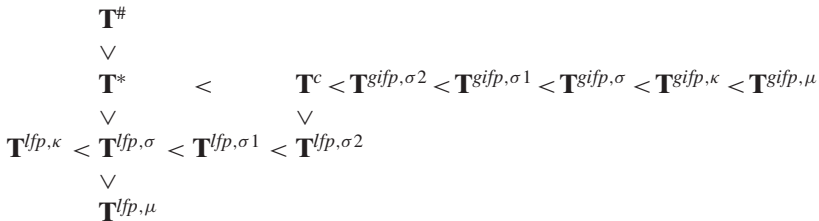
Proof Suppose that $\mathbf{T} \leq_3 \mathbf{T}'$ and $\mathbf{T}' \leq_3 \mathbf{T}''$, and that M is an $(S - Y)$ -neutral ground model where $Y \subseteq \{A : A \in \mathbf{V}_M \text{ or } \neg A \in \mathbf{V}_M\}$ and where $\mathbf{V}_M = \{A \in S : A \text{ is valid in the ground model } M \text{ according to } \mathbf{T}\}$. Let $\mathbf{V}'_M = \{A \in S : A \text{ is valid in the ground model } M \text{ according to } \mathbf{T}'\}$. Note that $S = \{A : A \in \mathbf{V}'_M \text{ or } \neg A \in \mathbf{V}'_M\}$, since $\mathbf{T} \leq_3 \mathbf{T}'$. So $Y \subseteq \{A : A \in \mathbf{V}'_M \text{ or } \neg A \in \mathbf{V}'_M\}$. So \mathbf{T}'' dictates that truth behaves like a classical concept in M , as desired. \square

Lemma 4.20 (1) If $\mathbf{T} \leq_3 \mathbf{T}'$ and $\mathbf{T}' \leq_2 \mathbf{T}''$ then $\mathbf{T} \leq_3 \mathbf{T}''$. (2) If $\mathbf{T} \leq_3 \mathbf{T}'$ then $\mathbf{T} \leq_2 \mathbf{T}'$. (3) If $\mathbf{T} \leq_1 \mathbf{T}'$ and $\mathbf{T}' \leq_3 \mathbf{T}''$ then $\mathbf{T} \leq_3 \mathbf{T}''$.

Proof (1) follows immediately from the definitions. For (2), suppose that $\mathbf{T} \leq_3 \mathbf{T}'$ and that \mathbf{T} dictates that truth behaves like a classical concept in M . Then M is $(S - S)$ -neutral where $S \subseteq \{A : A \in \mathbf{V}_M \text{ or } \neg A \in \mathbf{V}_M\}$. So \mathbf{T}' dictates that truth behaves like a classical concept in M , since $\mathbf{T} \leq_3 \mathbf{T}'$. For (3), assume that $\mathbf{T} \leq_1 \mathbf{T}'$ and $\mathbf{T}' \leq_3 \mathbf{T}''$ and that M is $(S - Y)$ -neutral where $Y \subseteq \{A : A \in \mathbf{V}_M \text{ or } \neg A \in \mathbf{V}_M\}$. Since $\mathbf{T} \leq_1 \mathbf{T}'$, M is $(S - Y)$ -neutral where $Y \subseteq \{A : A \in \mathbf{V}'_M \text{ or } \neg A \in \mathbf{V}'_M\}$. So, since $\mathbf{T}' \leq_3 \mathbf{T}''$, \mathbf{T}'' dictates that truth behaves like a classical concept in M , as desired. \square

Theorem 4.21

(1) The relation $<_3$ behaves as in the following diagram, i.e., $<_3$ is the smallest transitive relation satisfying the conditions given the diagram. Since \leq_3 is not reflexive, we need parts (2) and (3) to completely determine \leq_3 . The subscripted 3 has been dropped from the diagram:



- (2) (i) $\mathbf{T}^* \leq_3 \mathbf{T}^*$, (ii) $\mathbf{T}^c \leq_3 \mathbf{T}^c$, (iii) $\mathbf{T}^{lfp,\sigma 2} \leq_3 \mathbf{T}^{lfp,\sigma 2}$, and (iv) $\mathbf{T}^{gfp,\rho} \leq_3 \mathbf{T}^{gfp,\rho}$ for $\rho = \mu, \kappa, \sigma, \sigma 1$ or $\sigma 2$.
 (3) $\mathbf{T}^\# \not\leq_3 \mathbf{T}^\#$ and $\mathbf{T}^{lfp,\rho} \not\leq_3 \mathbf{T}^{lfp,\rho}$ for $\rho = \mu, \kappa, \sigma$ or $\sigma 1$.

Proof The proofs of (2) and (3) are tricky and left until Section 5. Given (2) and (3), and Lemma 4.20, much of the information contained in (1) can

be straightforwardly proved. First, every claim of the form $\mathbf{T} \not\leq_2 \mathbf{T}'$ given in Theorem 4.5 implies, given Lemma 4.20 (2), that $\mathbf{T} \not\leq_3 \mathbf{T}'$. Furthermore, the facts that $\mathbf{T}^{lfp,\mu} \not\leq_3 \mathbf{T}^{lfp,\kappa}$ and that $\mathbf{T}^{lfp,\kappa} \not\leq_3 \mathbf{T}^{lfp,\mu}$ follow from the fact that neither $lfp(\mu_M)$ nor $lfp(\kappa_M)$ is classical in any ground model M , even which the ground model is S -neutral, as shown in the proof of Theorem 4.5. The fact that $\mathbf{T}^* \leq_3 \mathbf{T}^c \leq_3 \mathbf{T}^{gifp,\sigma^2} \leq_3 \mathbf{T}^{gifp,\sigma^1} \leq_3 \mathbf{T}^{gifp,\sigma} \leq_3 \mathbf{T}^{gifp,\kappa} \leq_3 \mathbf{T}^{gifp,\mu}$ follows from the fact that $\mathbf{T}^* \leq_2 \mathbf{T}^c \leq_2 \mathbf{T}^{gifp,\sigma^2} \leq_2 \mathbf{T}^{gifp,\sigma^1} \leq_2 \mathbf{T}^{gifp,\sigma} \leq_2 \mathbf{T}^{gifp,\kappa} \leq_2 \mathbf{T}^{gifp,\mu}$ and from Theorem 4.21 (2) and Lemma 4.20 (1). Similarly for the fact that $\mathbf{T}^{lfp,\sigma^2} \leq_3 \mathbf{T}^c$. The fact that $\mathbf{T}^{lfp,\sigma^1} \leq_3 \mathbf{T}^{lfp,\sigma^2}$ follows from the fact that $\mathbf{T}^{lfp,\sigma^1} \leq_1 \mathbf{T}^{lfp,\sigma^2}$ (Theorem 4.5) and that $\mathbf{T}^{lfp,\sigma^2} \leq_3 \mathbf{T}^{lfp,\sigma^2}$ (Theorem 4.21 (2)) and from Lemma 4.20 (3).

So, for Theorem 4.21, it remains to show (2) and (3), as well as $\mathbf{T}^{lfp,\mu} \leq_3 \mathbf{T}^{lfp,\sigma} \leq_3 \mathbf{T}^{lfp,\sigma^1}$ and $\mathbf{T}^{lfp,\kappa} \leq_3 \mathbf{T}^{lfp,\sigma}$. For (2) and (3), see Section 5. For the rest, see Corollary 4.28. □

Remark 4.22 The positive part of Theorem 4.21 generalizes Gupta and Belnap’s Theorems 4.12 (ii), (iii) and (iv) and 4.15, stated above. The negative parts generalize the negative answer to Gupta and Belnap’s Question 4.16, asked above.

The fact that $\mathbf{T}^{lfp,\sigma} \not\leq_3 \mathbf{T}^{lfp,\sigma}$ means that the following conjecture is false: If the ground model M is $(S - Y)$ -neutral and $Y \subseteq \{A : lfp(\sigma_M)(A) = \mathbf{t} \text{ or } lfp(\sigma_M)(A) = \mathbf{f}\}$, then $lfp(\sigma_M)(A)$ is classical. Similarly for $\sigma 1$. But we have something almost as good in Theorem 4.24. First, a definition.

Definition 4.23 ([3], Definition 6A.2) The *degree* of a term or formula X of \mathcal{L}^+ , denoted $deg(X)$, is defined as follows. (i) If X is a variable or nonquote name, then $deg(X) = 0 = deg(\perp)$. (ii) If A is a sentence of degree n , then $deg(\ulcorner A \urcorner) = n + 1$. (iii) If t_1, \dots, t_n are terms of degrees i_1, \dots, i_n , respectively, and if F is an n -place function symbol or relation symbol, then $deg(Ft_1 \dots t_n) = \max(i_1, \dots, i_n)$. (iv) If x is a variable, and A and B are formulas of degrees m and n respectively, then $deg(\forall x A) = deg(\exists x A) = deg(\neg A) = m$ and $deg(A \ \& \ B) = deg(A \vee B) = \max(m, n)$ (and similarly if the language has other connectives, such as \supset or \equiv).

Theorem 4.24 (Proviso Theorem) *Let $\rho = \sigma$ or $\sigma 1$. If the ground model M is $(S - Y)$ -neutral and $Y \subseteq \{A : lfp(\rho_M)(A) = \mathbf{t} \text{ or } lfp(\rho_M)(A) = \mathbf{f}\}$, then $lfp(\rho_M)(A)$ is classical, subject to the following proviso: for every $n \geq 0$, the following two sets have the same cardinality as S :*

$$\begin{aligned} & \{B \in S - Y : deg(B) > n \ \& \ lfp(\rho_M)(B) = \mathbf{t}\}, \text{ and} \\ & \{B \in S - Y : deg(B) > n \ \& \ lfp(\rho_M)(B) = \mathbf{f}\}. \end{aligned}$$

Proof See Section 5, below. □

Remark 4.25 If S is countable, then the proviso in Theorem 4.24 is equivalent to the following: for every $n \geq 0$, there is a sentence $B \in S - Y$ of degree $> n$ such that $lfp(\rho_M)(B) = \mathbf{t}$, and a sentence $B \in S - Y$ of degree $> n$ such that $lfp(\rho_M)(B) = \mathbf{f}$.

Corollary 4.26 *If the ground model M is X -neutral, where X contains all sentences that have occurrences of \mathbf{T} , then the following theories dictate that truth behaves like a classical concept in M : $\mathbf{T}^{lfp,\sigma}$, $\mathbf{T}^{lfp,\sigma^1}$, $\mathbf{T}^{lfp,\sigma^2}$, $\mathbf{T}^{gifp,\mu}$, $\mathbf{T}^{gifp,\kappa}$, $\mathbf{T}^{gifp,\sigma}$, $\mathbf{T}^{gifp,\sigma^1}$, $\mathbf{T}^{gifp,\sigma^2}$, \mathbf{T}^* , $\mathbf{T}^\#$, \mathbf{T}^c . In particular, if the ground model M is S -neutral, then all those theories dictate that truth behaves like a classical concept in M .*

Proof We prove the result for $\mathbf{T}^{lfp,\sigma}$. For the other theories $\mathbf{T}^{lfp,\sigma^1}$, $\mathbf{T}^{lfp,\sigma^2}$, \mathbf{T}^* , $\mathbf{T}^\#$, \mathbf{T}^c and the $\mathbf{T}^{gifp,\rho}$, the result follows from this and the positive part of Theorem 4.5, which we have already proved.

Assume that the ground model M is X -neutral, where X contains all sentences that have occurrences of \mathbf{T} . Let $Y = \{A \in S : \mathbf{T}$ does not occur in $A\}$. So M is $(S - Y)$ -neutral and $Y \subseteq \{A : lfp(\sigma_M)(A) = \mathbf{t} \text{ or } lfp(\sigma_M)(A) = \mathbf{f}\}$. Also, we claim that the proviso in Theorem 4.24 is satisfied for $\rho = \sigma$. In particular, for any sentence A , define the sentences $\mathbf{T}^0(A) = A$ and $\mathbf{T}^{n+1}(A) = \mathbf{T} \cdot \mathbf{T}^n(A)$. Note that, for any $n \geq 0$ and any $A \in S$, we have,

$$\begin{aligned} \mathbf{T}^{n+1}(A \vee \neg A) &\in \{B \in S - Y : deg(B) > n \ \& \ lfp(\sigma_M)(B) = \mathbf{t}\}, \text{ and} \\ \mathbf{T}^{n+1}(A \ \& \ \neg A) &\in \{B \in S - Y : deg(B) > n \ \& \ lfp(\sigma_M)(B) = \mathbf{f}\}. \end{aligned}$$

This suffices for the proviso. □

Remark 4.27 Theorem 4.24 generalizes Gupta and Belnap’s Theorem 4.13 (i), stated above.

Corollary 4.28 $\mathbf{T}^{lfp,\kappa} \leq_3 \mathbf{T}^{lfp,\sigma} \leq_3 \mathbf{T}^{lfp,\sigma^1}$ and $\mathbf{T}^{lfp,\mu} \leq_3 \mathbf{T}^{lfp,\sigma}$.

Proof To see that $\mathbf{T}^{lfp,\sigma} \leq_3 \mathbf{T}^{lfp,\sigma^1}$, suppose that the ground model M is $(S - Y)$ -neutral and that $Y \subseteq \{A \in S : lfp(\sigma_M)(A) = \mathbf{t} \text{ or } lfp(\sigma_M)(A) = \mathbf{f}\}$. If $\mathbf{T}^{lfp,\sigma}$ dictates that truth behaves like a classical concept in M , then so does $\mathbf{T}^{lfp,\sigma^1}$, since $lfp(\sigma_M) \leq lfp(\sigma \ 1_M)$ (see the proof of Theorem 4.2). So suppose that $\mathbf{T}^{lfp,\sigma}$ does not dictate that truth behaves like a classical concept in M . First notice that $Y \subseteq \{A \in S : lfp(\sigma \ 1_M)(A) = \mathbf{t} \text{ or } lfp(\sigma \ 1_M)(A) = \mathbf{f}\}$.

Also, we claim that the proviso in Theorem 4.24 is satisfied for $\rho = \sigma \ 1$. To see this, choose some sentence C such that $lfp(\sigma_M)(C) = \mathbf{n}$. And let C^* be the sentence $(\mathbf{T} \cdot C \ \& \ \mathbf{T} \cdot \neg C)$. Note that $lfp(\sigma_M)(\neg C) = \mathbf{n}$. So there are classical hypotheses $h, h' \geq lfp(\sigma_M)$ such that $h(C) = h(\neg C) = \mathbf{t}$ and $h'(C) = h'(\neg C) = \mathbf{f}$. So $Val_{M+h,\tau}(\mathbf{T} \cdot C \ \& \ \mathbf{T} \cdot \neg C) = \mathbf{t}$ and $Val_{M+h',\tau}(\mathbf{T} \cdot C \ \& \ \mathbf{T} \cdot \neg C) = \mathbf{f}$.

So $lfp(\sigma_M)(C^*) = \mathbf{n}$. So $lfp(\sigma_M)(\neg C^*) = \mathbf{n}$. So for any $n \geq 0$, both $lfp(\sigma_M)(\mathbf{T}^n(C^*)) = \mathbf{n}$ and $lfp(\sigma_M)(\mathbf{T}^n(\neg C^*)) = \mathbf{n}$. So for any $n \geq 0$ and any $A \in S$, both

$$lfp(\sigma_M)(\mathbf{T}^n(C^*) \ \& \ 'A' = 'A') = \mathbf{n} \text{ and}$$

$$lfp(\sigma_M)(\mathbf{T}^n(\neg C^*) \vee 'A' \neq 'A') = \mathbf{n}.$$

So for any $n \geq 0$ and any $A \in S$, both

$$(\mathbf{T}^n(C^*) \ \& \ 'A' = 'A') \in S - Y \text{ and}$$

$$(\mathbf{T}^n(\neg C^*) \vee 'A' \neq 'A') \in S - Y.$$

So for any $n \geq 0$ and any $A \in S$, both

$$(\mathbf{T}^n(C^*) \ \& \ 'A' = 'A') \in \{B \in S - Y : deg(B) > n \ \& \ lfp(\sigma_{1_M})(B) = \mathbf{t}\}, \text{ and}$$

$$(\mathbf{T}^n(\neg C^*) \vee 'A' \neq 'A') \in \{B \in S - Y : deg(B) > n \ \& \ lfp(\sigma_{1_M})(B) = \mathbf{f}\}.$$

This suffices for the proviso.

The proof that $\mathbf{T}^{lfp,\kappa} \leq_3 \mathbf{T}^{lfp,\sigma}$ is similar. Suppose that the ground model M is $(S - Y)$ -neutral where $Y \subseteq \{A \in S : lfp(\kappa_M)(A) = \mathbf{t} \text{ or } lfp(\kappa_M)(A) = \mathbf{f}\}$. First notice that $Y \subseteq \{A \in S : lfp(\sigma_M)(A) = \mathbf{t} \text{ or } lfp(\sigma_M)(A) = \mathbf{f}\}$. Also, we claim that the proviso in Theorem 4.24 is satisfied for $\rho = \sigma$. To see this, let C be the sentence $\forall x(\mathbf{T}x \vee \neg \mathbf{T}x)$. Note that, for any $n \geq 0$ and any $A \in S$, both

$$lfp(\kappa_M)(\mathbf{T}^{n+1}(C) \ \& \ 'A' = 'A') = \mathbf{n} \text{ and}$$

$$lfp(\kappa_M)(\mathbf{T}^{n+1}(\neg C) \vee 'A' \neq 'A') = \mathbf{n}.$$

So, for any $n \geq 0$ and any $A \in S$, both

$$(\mathbf{T}^{n+1}(C) \ \& \ 'A' = 'A') \in S - Y \text{ and}$$

$$(\mathbf{T}^{n+1}(\neg C) \vee 'A' \neq 'A') \in S - Y.$$

So for any $n \geq 0$ and any $A \in S$, both

$$(\mathbf{T}^{n+1}(C) \ \& \ 'A' = 'A') \in \{B \in S - Y : deg(B) > n \ \& \ lfp(\sigma_M)(B) = \mathbf{t}\}, \text{ and}$$

$$(\mathbf{T}^{n+1}(\neg C) \vee 'A' \neq 'A') \in \{B \in S - Y : deg(B) > n \ \& \ lfp(\sigma_M)(B) = \mathbf{f}\}.$$

This suffices for the proviso.

Similarly, $\mathbf{T}^{lfp,\mu} \leq_3 \mathbf{T}^{lfp,\sigma}$. □

5 Proofs and Counterexamples

Each of our main theorems, Theorems 4.2, 4.5 and 4.21, makes positive claims of the form $\mathbf{T} \leq_n \mathbf{T}'$ and negative claims of the form $\mathbf{T} \not\leq_n \mathbf{T}'$. We also want to

show Theorem 4.24 (the Proviso Theorem). Given the work already done in Section 4, it suffices to show Theorem 4.21 (2) and (3), to show Theorem 4.24, and to show the negative claims of Theorems 4.2 and 4.5.

We begin with some preliminary notions. Then we prove our *Major Lemma* (Lemma 5.4) and *Major Corollary* (Corollary 5.5), which we will use to help establish our results from Section 4. Before that we will use the Major Corollary to give a simplified proof of Gupta and Belnap’s *Main Lemma* from [3] (Lemma 5.6), the lemma that they use to study the conditions under which a model is Thomason: our new proof avoids their double transfinite induction, and their consideration, at one point, of six cases and subcases.

Definition 5.1 Suppose that $M = \langle D, I \rangle$ and $M' = \langle D', I' \rangle$ are models of a first-order language \mathcal{L} , that N is a set of names of \mathcal{L} , and that $\Psi : D \rightarrow D'$ is a bijection. The bijection Ψ is an *N-restricted isomorphism from M to M'* iff $I(R)(d_1, \dots, d_n) = I'(R)(\Psi(d_1), \dots, \Psi(d_n))$ for every n -place relation symbol R and every $d_1, \dots, d_n \in D$; $\Psi(I(f)(d_1, \dots, d_n)) = I'(f)(\Psi(d_1), \dots, \Psi(d_n))$ for every n -place function symbols ($n \geq 1$) and every $d_1, \dots, d_n \in D$; and $\Psi(I(c)) = I'(c)$ for every $c \in N$.

Lemma 5.2 *Suppose that M and M' are models of a first-order language L, that N is a set of names in L, and that Psi is an N-restricted isomorphism from M to M'. Suppose that rho = tau, mu, kappa or sigma. Suppose that every name occurring in the sentence A is in N. Then Val_{M,rho}(A) = Val_{M',rho}(A).*

Definition 5.3 Suppose that $M = \langle D, I \rangle$ is a ground model for a ground language \mathcal{L} , and that $Y \subseteq S =$ the set of sentences of \mathcal{L}^+ . We say that a hypothesis h is *superstrongly consistent* if h is strongly consistent and $h(d) = \mathbf{f}$ for every $d \in D - S$. We say that $h =_Y h'$ iff $h(A) = h'(A)$ for every $A \in Y$. If n is a natural number, we say that $h =_n h'$ iff $h(A) = h'(A)$ for every sentence A of \mathcal{L}^+ of degree $< n$. Note that $h =_0 h'$ for any h and h' . If h is a classical hypothesis, define $\tau_M^0(h) = h$ and $\tau_0^{n+1}(h) = \tau_M(\tau_M^n(h))$. Define $\tau_M^\omega(h) : D \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{n}\}$ as follows:

$$\tau_M^\omega(h)(d) = \begin{cases} \mathbf{t}, & \text{if } (\exists m \geq 0)(\forall n \geq m)(\tau_M^n(h)(d) = \mathbf{t}) \\ \mathbf{f}, & \text{if } (\exists m \geq 0)(\forall n \geq m)(\tau_M^n(h)(d) = \mathbf{f}) \\ \mathbf{n}, & \text{otherwise.} \end{cases}$$

Note that if h is classical, then $\tau_M^n(h)$ is always classical, but $\tau_M^\omega(h)$ might not be.

Lemma 5.4 (Major Lemma) *Suppose that M = <D, I> is a ground model for a ground language L, and that Y subseteq S = the set of sentences of L+. Suppose that M is (S - Y)-neutral, and that h and h' are superstrongly consistent classical hypotheses, with h =_n h' and h =_Y h'. Then tau_M(h) =_{n+1} tau_M(h').*

Proof Make the assumptions in the statement of the lemma. Let $Y^* = \{A \in S : h(A) = h'(A)\}$. Note: (1) $Y \subseteq Y^*$, (2) $h =_{Y^*} h'$, (3) if $\text{deg}(A) < n$ then $A \in Y^*$, and (4) M is $(S - Y^*)$ -neutral. Define

$$U =_{\text{df}} \{A \in S - Y^* : h(A) = \mathbf{t}\}$$

$$V =_{\text{df}} \{A \in S - Y^* : h(A) = \mathbf{f}\}$$

Note that $U \cup V = S - Y^*$. Given the definition of Y^* , if $A \in S - Y^*$, then $h(A) \neq h'(A)$. Thus,

$$U = \{A \in S - Y^* : h'(A) = \mathbf{f}\}$$

$$V = \{A \in S - Y^* : h'(A) = \mathbf{t}\}$$

Now notice that, for every $A \in S - Y^*$, we have $A \in U$ iff $\neg A \in V$ and $A \in V$ iff $\neg A \in U$. So U and V have the same cardinality (which might be 0, or some infinite cardinality).

Let Φ be a bijection from U onto V . Define a function $\Psi : D \rightarrow D$ as follows:

$$\Psi(A) = \begin{cases} A, & \text{if } A \in Y^* \\ \Phi(A), & \text{if } A \in U \\ \Phi^{-1}(A), & \text{if } A \in V \end{cases}$$

$$\Psi(d) = d, \text{ if } d \in (D - S)$$

Note that Ψ is an N -restricted isomorphism from $M + h$ to $M + h'$, where N is the set of names of degree $\leq n$. So $\text{Val}_{M+h, \tau}(A) = \text{Val}_{M+h', \tau}(A)$, for every sentence A of degree $< n + 1$. So $\tau_M(h) =_{n+1} \tau_M(h')$, as desired. \square

Corollary 5.5 (Major Corollary) *Suppose that the ground model $M = \langle D, I \rangle$ is $(S - Y)$ -neutral, where $Y \subseteq S$. Suppose that h and h' are classical hypotheses such that $\tau_M^n(h) =_Y \tau_M^{n+1}(h) =_Y \tau_M^n(h') =_Y \tau_M^{n+1}(h')$, for every $n \geq 0$. Then $\tau_M^\omega(h) = \tau_M^\omega(h')$ is a classical fixed point of τ_M .*

Proof Make the assumptions in the statement of the corollary. Note that, for $n \geq 0$, the hypotheses $\tau_M^{n+1}(h)$ and $\tau_M^{n+1}(h')$ are superstrongly consistent. By induction, we can show that $\tau_M^{n+1}(h) =_n \tau_M^{n+2}(h) =_n \tau_M^{n+1}(h') =_n \tau_M^{n+2}(h')$, for every $n \geq 0$. The base case is vacuously true. The induction step is simply an application of the Major Lemma. But from this it follows that $\tau_M^\omega(h) = \tau_M^\omega(h')$ and $\tau_M^\omega(h)$ is classical.

Note also that $\tau_M^\omega(h)$ is strongly consistent. To see this, suppose not: suppose that the set $\{A \in S : \tau_M^\omega(h)(A) = \mathbf{t}\} \cup \{\neg A \in S : \tau_M^\omega(h)(A) = \mathbf{f}\}$ is inconsistent. Then there are sentence A_1, \dots, A_k and B_1, \dots, B_m with $\tau_M^\omega(h)(A_i) = \mathbf{t}$ and $\tau_M^\omega(h)(B_j) = \mathbf{f}$, and such that $\{A_1, \dots, A_k, \neg B_1, \dots, \neg B_m\}$ is inconsistent. Let $n = \max(\text{deg}(A_1), \dots, \text{deg}(A_k), \text{deg}(B_1), \dots, \text{deg}(B_m))$. Note that $\tau_M^\omega(h) =_{n+1}$

$\tau_M^{n+2}(h)$. So $\tau_M^{n+2}(h)(A_i) = \mathbf{t}$ and $\tau_M^{n+2}(h)(B_i) = \mathbf{f}$. So $\tau_M^{n+2}(h)$ is not strongly consistent. But this cannot be, since the application of τ_M to any classical hypothesis delivers a strongly consistent, indeed a superstrongly consistent, classical hypothesis.

Not only is $\tau_M^\omega(h)$ strongly consistent; it is superstrongly consistent. This follows from the fact that $\tau_M^{n+1}(h)(d) = \mathbf{f}$, for every $n \geq 0$ and every $d \in D - S$.

We must still show that $\tau_M^\omega(h)$ is a fixed point of τ_M . Note that both $\tau_M^\omega(h) =_Y \tau_M^{n+1}(h)$ and $\tau_M^\omega(h) =_n \tau_M^{n+1}(h)$, for every $n \geq 0$. So, by the Major Lemma, $\tau_M(\tau_M^\omega(h)) =_{n+1} \tau_M^{n+2}(h)$, for every $n \geq 0$. So $\tau_M(\tau_M^\omega(h)) =_{n+1} \tau_M^\omega(h)$, for every $n \geq 0$. So $\tau_M(\tau_M^\omega(h)) = \tau_M^\omega(h)$, as desired. \square

Lemma 5.6 (Gupta and Belnap’s Main Lemma, [3], Lemma 6A.4) *Let $M = \langle D, I \rangle$ be X -neutral ($X \subseteq D$). Let \mathcal{S} and \mathcal{S}' be τ_M -sequences, and let Y be the set of those sentences that are either stably \mathbf{t} in both \mathcal{S} and \mathcal{S}' or stably \mathbf{f} in both. If $(S - Y) \subseteq X$, then there is some ordinal α such that for all $\beta \geq \alpha$, $\mathcal{S}_\alpha = \mathcal{S}'_\beta$.*

Proof This proof differs from Gupta and Belnap’s. Choose an ordinal γ such that, by the γ^{th} stage both in \mathcal{S} and \mathcal{S}' , all of the sentences in Y have stabilized: for every $A \in Y$ and every $\beta \geq \gamma$, we have $\mathcal{S}_\beta(A) = \mathcal{S}'_\beta(A) = \mathcal{S}_\gamma(A) = \mathcal{S}'_\gamma(A)$. In other words, for every $\beta \geq \gamma$, we have $\mathcal{S}_\beta =_Y \mathcal{S}'_\beta =_Y \mathcal{S}_\gamma =_Y \mathcal{S}'_\gamma$. By our Major Corollary, $\tau_M^\omega(\mathcal{S}_\gamma) = \tau_M^\omega(\mathcal{S}'_\gamma)$ is classical and is a fixed point of τ_M . But notice that, since $\tau_M^\omega(\mathcal{S}_\gamma) = \tau_M^\omega(\mathcal{S}'_\gamma)$ is classical, we have $\mathcal{S}_{\gamma+\omega} = \tau_M^\omega(\mathcal{S}_\gamma)$ and $\mathcal{S}'_{\gamma+\omega} = \tau_M^\omega(\mathcal{S}'_\gamma)$ by the limit rule for τ_M -sequences. Let $\alpha = \gamma + \omega$. Since $\mathcal{S}_\alpha = \mathcal{S}'_\alpha$ is a fixed point of τ_M , we conclude that $\mathcal{S}_\alpha = \mathcal{S}'_\beta$ for all $\beta \geq \alpha$, as desired. \square

Now we can prove our positive results from Section 4: these are Theorem 4.21 (2), and Theorem 4.24.

Proof of Theorem 4.21 (2)

- (i) (The proof of (i) is from [3].) Suppose that M is an $(S - Y)$ -neutral model and that $Y \subseteq \{A \in S : A \in \mathbf{V}_M^* \text{ or } \neg A \in \mathbf{V}_M^*\}$. We want to show that \mathbf{T}^* dictates that truth behaves like a classical concept in M . For this it suffices to show that all τ_M -sequences culminate in one and the same fixed point. Choose any two τ_M -sequences \mathcal{S} and \mathcal{S}' . Let $X = (S - Y)$, and let Y^* be the set of those sentences that are either stably \mathbf{t} in both \mathcal{S} and \mathcal{S}' or stably \mathbf{f} in both. Clearly $(S - Y^*) \subseteq X$. So, by Gupta and Belnap’s Main Lemma (our Lemma 5.6, above), there is some ordinal α such that $\mathcal{S}_\alpha = \mathcal{S}'_\beta$ for all $\beta \geq \alpha$. It follows that \mathcal{S}_α is a fixed point in which both \mathcal{S} and \mathcal{S}' culminate.
- (ii) is proved analogously to (i), since it suffices to show that if M is an $(S - Y)$ -neutral model where $Y \subseteq \{A \in S : A \in \mathbf{V}_M^c \text{ or } \neg A \in \mathbf{V}_M^c\}$, then all

maximally consistent τ_M -sequences culminate in one and the same fixed point.

- (iii) Suppose that M is an $(S - Y)$ -neutral model and that $Y \subseteq \{A \in S : lfp(\sigma_{2_M})(A) = \mathbf{t} \text{ or } lfp(\sigma_{2_M})(A) = \mathbf{f}\}$. To show that $h = lfp(\sigma_2)$ is classical, suppose not. Let C be a sentence of the least possible degree, say k , such that $h(C) = \mathbf{n}$. Note that $C \notin Y$. We will get a contradiction by showing that $h(C) = \mathbf{t}$ or \mathbf{f} . Recall the definition of $\sigma_{2_M}(h)(A)$ for sentences A :

$$\sigma_{2_M}(h)(A) = \begin{cases} \mathbf{t}, & \text{if } \tau_M(h')(A) = \mathbf{t} \text{ for every strongly consistent classical} \\ & h' \geq h \\ \mathbf{f}, & \text{if } \tau_M(h')(A) = \mathbf{f} \text{ for every strongly consistent classical} \\ & h' \geq h \\ \mathbf{n}, & \text{otherwise, for sentences } A \in S. \end{cases}$$

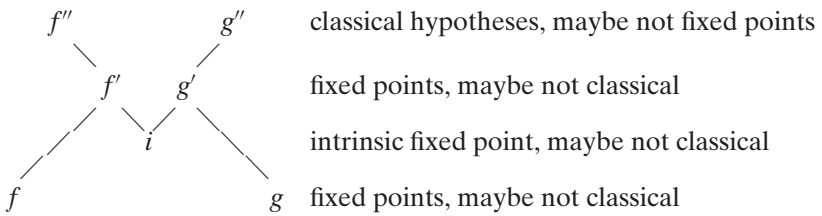
To show that $h(C) = \mathbf{t}$ or \mathbf{f} , it suffices to show that $\sigma_{2_M}(h)(C) = \mathbf{t}$ or \mathbf{f} , since h is a fixed point of σ_{2_M} . For this, it suffices to show that $\tau_M(h')(C) = \tau_M(h'')(C)$ for any classical strongly consistent hypotheses $h', h'' \geq h$. Choose such hypotheses h' and h'' . Note that h' and h'' are *superstrongly* consistent, since $h(d) = \mathbf{f}$ for every $d \in D - S$. Note that $h' =_k h''$ since $h(A) = \mathbf{t}$ or \mathbf{f} , for any sentence A of degree $< k$. Note also that $h' =_Y h''$. So by our Major Lemma 5.4, $\tau_M(h') =_{k+1} \tau_M(h'')$. So $\tau_M(h')(C) = \tau_M(h'')(C)$, as desired.

- (iv) We will show something more general. Fix a ground model M . If \mathcal{F} is a partial function on the set of hypotheses, we say that \mathcal{F} is *normal* iff \mathcal{F} satisfies the following conditions: \mathcal{F} is monotone; if h is classical and \mathcal{F} is defined on h then $\mathcal{F}(h) = \tau_M(h)$; for every fixed point h of \mathcal{F} , there is a classical hypothesis $h' \geq h$ such that \mathcal{F} is defined on h' ; if \mathcal{F} is defined on the classical hypothesis h , then \mathcal{F} is defined on $\tau_M(h)$; and \mathcal{F} is defined on every fixed point of τ_M . Note that $\mu_M, \kappa_M, \sigma_M, \sigma 1_M$ and $\sigma 2_M$ are all normal.

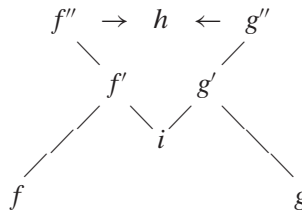
Suppose that \mathcal{F} is a normal partial function on hypotheses, and that i is an intrinsic fixed point of \mathcal{F} . Suppose that M is $(S - Y)$ -neutral where $i(A) = \mathbf{t}$ or $i(A) = \mathbf{f}$ for every sentence $A \in Y$. We will show that $gifp(\mathcal{F})$ is classical. This will suffice for our claim that $\mathbf{T}^{gifp,\rho} \leq_3 \mathbf{T}^{gifp,\rho}$ for $\rho = \mu, \kappa, \sigma, \sigma 1$ or $\sigma 2$.

To show that $gifp(\mathcal{F})$ is classical, it will suffice to show that \mathcal{F} has a greatest fixed point, which is classical: any classical greatest fixed point is also the greatest intrinsic fixed point. For this, it suffices to show that for any fixed points f and g , there is a classical fixed point $h \geq f, g$. So choose any fixed points f and g . Since i is intrinsic, there are fixed points f' and g' such that $f, i \leq f'$ and $g, i \leq g'$. Choose classical hypotheses,

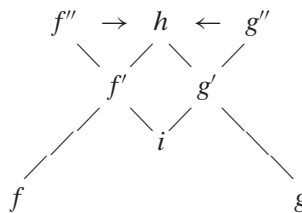
not necessarily fixed points, $f'' \geq f'$ and $g'' \geq g'$, so that \mathcal{F} is defined on both f'' and g'' . Here is a picture:



Note: $\tau_M^n(f'') = \mathcal{F}^n(f'') \geq \mathcal{F}^n(f') = f' \geq i$ and $\tau_M^n(g'') = \mathcal{F}^n(g'') \geq \mathcal{F}^n(g') = g' \geq i$, for every $n \geq 0$. Recall that $Y \subseteq \{A : i(A) = \mathbf{t} \text{ or } i(A) = \mathbf{f}\}$. So $\tau_M^n(f'') =_Y i =_Y \tau_M^n(g'')$, for every $n \geq 0$. So $\tau_M^n(f'') =_Y \tau_M^{n+1}(f'') =_Y \tau_M^n(g'') =_Y \tau_M^{n+1}(g'')$, for every $n \geq 0$. Let $h = \tau_M^\omega(f'')$. By our Major Corollary 5.5, $h = \tau_M^\omega(f'') = \tau_M^\omega(g'')$ is a classical fixed point of τ_M . So h is a classical fixed point of \mathcal{F} . In the picture below, the arrow from f'' to h indicates that any revision sequence, indeed the *only* revision sequence, that begins with f'' culminates in h . Similarly for the arrow from g'' to h .



It now suffices to show that $h \geq f$ and $h \geq g$. For this, it suffices to show that $h \geq f'$ and $h \geq g'$. Note that if $f'(A) = \mathbf{t}$, then $\tau_M^n(f'')(A) = \mathbf{t}$ for every $n \geq 0$, since $\tau_M^n(f'') \geq f'$. So $h(A) = \tau_M^\omega(f'')(A) = \mathbf{t}$. We have just shown that $f'(A) = \mathbf{t} \Rightarrow h(A) = \mathbf{t}$, for every $A \in S$. Similarly, $f'(A) = \mathbf{f} \Rightarrow h(A) = \mathbf{f}$, for every $A \in S$. So $h \geq f'$. Similarly, $h \geq g'$, as desired. Our completed picture is as follows:



□

Proof of Theorem 4.24 Let $\rho = \sigma$ or $\sigma 1$, and make the assumptions in the statement of Theorem 4.24. In this proof, any parenthetical occurrence of ‘weakly consistent’ should be included if $\rho = \sigma 1$, and deleted if $\rho = \sigma$. For a reductio, suppose that $h_0 = lfp(\rho_M)$ is not classical.

Let C be a sentence of the least possible degree, say k , such that $h_0(C) = \mathbf{n}$. We will get a contradiction by showing that $h_0(C) = \mathbf{t}$ or \mathbf{f} . Recall that, for any sentence A ,

$$\rho_M(h_0)(A) = \begin{cases} \mathbf{t}, & \text{if } \tau_M(h)(A) = \mathbf{t} \text{ for each (weakly consistent) classical } h \geq h_0 \\ \mathbf{f}, & \text{if } \tau_M(h)(A) = \mathbf{f} \text{ for each (weakly consistent) classical } h \geq h_0 \\ \mathbf{n}, & \text{otherwise.} \end{cases}$$

Since $h_0 = \rho_M(h_0)$, in order to show that $h_0(C) = \mathbf{t}$ or \mathbf{f} , it will suffice to show that $\tau_M(h)(C) = \tau_M(h')(C)$, for any (weakly consistent) classical hypotheses $h, h' \geq h_0$. Choose such hypotheses h and h' . Note that $h =_k h'$, since $h_0(A) = \mathbf{t}$ or \mathbf{f} , for any sentence A of degree $< k$. Note also that $h =_Y h'$.

Define six sets:

$$\begin{aligned} U_0 &=_{\text{df}} \{A \in S - Y : \text{deg}(A) \geq k \ \& \ h_0(A) = \mathbf{t}\} \\ V_0 &=_{\text{df}} \{A \in S - Y : \text{deg}(A) \geq k \ \& \ h_0(A) = \mathbf{f}\} \\ U &=_{\text{df}} \{A \in S - Y : \text{deg}(A) \geq k \ \& \ h(A) = \mathbf{t}\} \\ V &=_{\text{df}} \{A \in S - Y : \text{deg}(A) \geq k \ \& \ h(A) = \mathbf{f}\} \\ U' &=_{\text{df}} \{A \in S - Y : \text{deg}(A) \geq k \ \& \ h'(A) = \mathbf{t}\} \\ V' &=_{\text{df}} \{A \in S - Y : \text{deg}(A) \geq k \ \& \ h'(A) = \mathbf{f}\} \end{aligned}$$

By the proviso, U_0 and V_0 have the same cardinality as S . Also, note that $U_0 \subseteq U \subseteq S$, so U has the same cardinality as S . Similarly, so do V, U' and V' .

Define a function $\Psi : D \rightarrow D$ by patching together the identity function on sentences of degree $< k$, the identity function on Y , the identity function on $D - S$, a bijection from U onto U' , and a bijection from V onto V' . Note that Ψ is an N -restricted isomorphism from $M + h$ to $M + h'$, where N is the set of names of degree $\leq k$. So $Val_{M+h,\tau}(A) = Val_{M+h',\tau}(A)$, for every sentence A of degree $< k + 1$. In particular, $Val_{M+h,\tau}(C) = Val_{M+h',\tau}(C)$. So $\tau_M(h)(C) = \tau_M(h')(C)$, as desired. □

What's left to prove is Theorem 4.21 (3), and the negative claims in Theorems 4.2 and 4.5. We do this with a series of counterexamples. We will bring it all together after presenting the examples.

Example 5.7 ([3], Example 6B.9) This example will show that $\mathbf{T}^\# \not\leq_3 \mathbf{T}^\#$. Consider a ground language \mathcal{L} with a one-place predicate G , and no other nonlogical vocabulary besides quote names. Let $A = \exists x(Gx \ \& \ \neg Tx)$ and let $Y = \{\mathbf{T}^n A : n \geq 0\}$. Let M be the ground model $\langle S, I \rangle$ where $I(G)(C) = \mathbf{t}$ iff $C \in Y$, for every $C \in S$. Note that every sentence in Y is nearly stably \mathbf{t} in every τ_M -sequence, though no sentence in Y is stably \mathbf{t} in any τ_M -sequence. So $C \in \mathbf{V}_M^\#$, for all $C \in Y$. So M is $(S - Y)$ -neutral where $Y \subseteq \{C : C \in \mathbf{V}_M^\# \text{ or } \neg C \in \mathbf{V}_M^\#\}$.

We will now show that there is a τ_M -sequence \mathcal{S} such that the sentence $B = \exists y \exists y(Gx \ \& \ Gy \ \& \ \neg Tx \ \& \ \neg Ty \ \& \ x \neq y)$ is neither nearly stably **t** in \mathcal{S} nor nearly stably **f** in \mathcal{S} . So $\mathbf{T}^\#$ does not dictate that truth behaves like a classical concept in M . Incidentally, this falsifies the claim in [3] that “all sentences are nearly stable in all τ -sequences for M ” (p. 214).

Define the sets $X_0 = Y$ and $X_{n+1} = Y - \{\mathbf{T}^n A\}$ for $n \geq 0$. Also define $Z_n = Y - \{\mathbf{T}^n A, \mathbf{T}^{n+1} A\}$. Then there is a τ_M -sequence \mathcal{S} such that, for each $C \in Y$, each limit ordinal λ and each $n \geq 0$,

$$\begin{aligned} \mathcal{S}_n(C) = \mathbf{t} & \text{ iff } C \in X_n, \text{ and} \\ \mathcal{S}_{\lambda+\omega^2+n}(C) = \mathbf{t} & \text{ iff } C \in Z_n, \text{ and} \\ \mathcal{S}_{\lambda+n}(C) = \mathbf{t} & \text{ iff } C \in X_n \text{ and } \lambda \text{ is a limit ordinal not of the form } \alpha + \omega^2. \end{aligned}$$

Note that $\mathcal{S}_{\lambda+\omega^2+n+1}(B) = \mathbf{t}$ and $\mathcal{S}_{\lambda+\omega+n+1}(B) = \mathbf{f}$, for every limit ordinal λ and every natural number n . So B is neither nearly stably **t**, nor nearly stably **f**, in \mathcal{S} .

Example 5.8 (Gupta) This example will show that $\mathbf{T}^\# \not\leq_2 \mathbf{T}^*$ and that $\mathbf{T}^\# \not\leq_2 \mathbf{T}^{gfp,\mu}$. Modify Example 5.7 as follows: Let Y be the smallest set containing each $\mathbf{T}^n A$ and such that if $C \in Y$ then $C \vee C \in Y$. Note that every sentence in Y is nearly stably **t** in every revision sequence, but no sentence in Y is stably **t** or stably **f** in any revision sequence. So τ_M has no classical fixed point. So neither \mathbf{T}^* nor $\mathbf{T}^{gfp,\mu}$ dictates that truth behaves like a classical concept in M . But it follows from Claim 2, below, that $\mathbf{T}^\#$ does dictate that truth behaves like a classical concept in M .

Notice that, for any classical hypothesis h and any $n \geq 0$, each of the following four sets is countably infinite:

$$\begin{aligned} \{C \in Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{t}\} \\ \{C \in Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{f}\} \\ \{C \notin Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{t}\} \\ \{C \notin Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{f}\} \end{aligned}$$

Claim 1 For any two classical hypotheses h and h' and any $n \geq 0$, $\tau_M^{n+2}(h) =_n \tau_M^{n+2}(h')$. We prove this by induction on n . The base case is vacuously true. For the inductive step, assume that $\tau_M^{n+2}(h) =_n \tau_M^{n+2}(h')$. To show that $\tau_M^{n+3}(h) =_{n+1} \tau_M^{n+3}(h')$, we construct an N -restricted isomorphism Ψ from $M + \tau_M^{n+2}(h)$ to

$M + \tau_M^{n+2}(h')$, where $N = \{‘C’ : deg(C) < n\}$. Define $U, U', V, V', W, W', X, X'$ as follows:

$$\begin{aligned}
 U &=_{\text{df}} \{C \in Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{t}\} \\
 U' &=_{\text{df}} \{C \in Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h')(C) = \mathbf{t}\} \\
 V &=_{\text{df}} \{C \in Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{f}\} \\
 V' &=_{\text{df}} \{C \in Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h')(C) = \mathbf{f}\} \\
 W &=_{\text{df}} \{C \in S - Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{t}\} \\
 W' &=_{\text{df}} \{C \in S - Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h')(C) = \mathbf{t}\} \\
 X &=_{\text{df}} \{C \in S - Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h)(C) = \mathbf{f}\} \\
 X' &=_{\text{df}} \{C \in S - Y : deg(C) \geq n \ \& \ \tau_M^{n+2}(h')(C) = \mathbf{f}\}
 \end{aligned}$$

Each of these sets is countably infinite. Define Ψ by patching together the identity function on the sentences of degree $< n$, and bijections from U onto U' , from V onto V' , from W onto W' , and from X onto X' .

Claim 2 For any sentence A of degree $< n$, either (i) $\tau_M^m(h)(A) = \mathbf{t}$ for every classical hypothesis h and every $m \geq n + 2$, or (ii) $\tau_M^m(h)(A) = \mathbf{f}$ for every classical hypothesis h and every $m \geq n + 2$. To see this, consider any classical hypotheses h and h' and any $m, m' \geq n + 2$. Note that if we apply Claim 1 to $\tau_M^{m-(n+2)}(h)$ and to $\tau_M^{m'-(n+2)}(h')$, we get $\tau_M^m(h) = \tau_M^{m'}(h')$.

From Claim 2, it follows that $\mathbf{T}^\#$ dictates that truth behaves like a classical concept in M . To see this, suppose that A is a sentence of \mathcal{L}^+ . By Claim 2, either (i) $\tau_M^m(h)(A) = \mathbf{t}$ for every classical hypothesis h and every $m \geq deg(A) + 3$, or (ii) $\tau_M^m(h)(A) = \mathbf{f}$ for every classical hypothesis h and every $m \geq deg(A) + 3$. Suppose (i). Note: for every τ_M -sequence \mathcal{S} , for every ordinal $\gamma \geq 0$, and for every $m \geq deg(A) + 3$, $\mathcal{S}_{\gamma+m}(A) = \mathbf{t}$. So A is nearly stably \mathbf{t} in every τ_M -sequence \mathcal{S} . So $A \in \mathbf{V}_M^\#$. Similarly, if (ii), then $\neg A \in \mathbf{V}_M^\#$. So, for every $A \in S$, either $A \in \mathbf{V}_M^\#$ or $\neg A \in \mathbf{V}_M^\#$. So $\mathbf{T}^\#$ dictates that truth behaves like a classical concept in M .

Example 5.9 This example presents a ground model M such that $lfp(\sigma 1_M)$ and $lfp(\sigma 2_M)$ are classical, and furthermore such that M is $(S - Y)$ -neutral where $Y \subseteq \{B : B \in \mathbf{V}_M^c \text{ or } \neg B \in \mathbf{V}_M^c\}$. On the negative side, we will see that neither \mathbf{T}^* nor $\mathbf{T}^\#$ dictates that truth behaves like a classical concept in M . This delivers the following results: $\mathbf{T}^{lfp,\sigma^1} \not\leq_2 \mathbf{T}^*$, $\mathbf{T}^{lfp,\sigma^1} \not\leq_2 \mathbf{T}^\#$, $\mathbf{T}^{lfp,\sigma^2} \not\leq_2 \mathbf{T}^*$, $\mathbf{T}^{lfp,\sigma^2} \not\leq_2 \mathbf{T}^\#$, $\mathbf{T}^c \not\leq_3 \mathbf{T}^*$, $\mathbf{T}^c \not\leq_3 \mathbf{T}^\#$. From these it follows—given Theorem 4.21 and Lemma 4.20—that $\mathbf{T}^c \not\leq_2 \mathbf{T}^*$ and $\mathbf{T}^c \not\leq_2 \mathbf{T}^\#$. The fact that $\mathbf{T}^c \not\leq_3 \mathbf{T}^*$ gives a negative answer to Gupta and Belnap’s Question 4.16, above.

Consider a ground language \mathcal{L} with a one-place predicate G , a nonquote name b , and no other nonlogical vocabulary besides quote names. For any formula B and $n \geq 0$, we define $-^n B$ as B when n is even and as $\neg B$ when n is odd. Let $A = \mathbf{T}'\mathbf{T}b'$ & $\mathbf{T}'\neg\mathbf{T}b'$. Let $Z = \{\mathbf{T}^n A : n \geq 0\}$. Let $Y = Z \cup \{\exists x(Gx \ \& \ \mathbf{T}x) \ \& \ \neg\mathbf{T}b\}$. Let M be the ground model $\langle S, I \rangle$, where $I(b) = \exists x(Gx \ \& \ \mathbf{T}x) \ \& \ \neg\mathbf{T}b$ and where $I(G)(C) = \mathbf{t}$ iff $C \in Z$, for every $C \in S$. Note that M is $(S - Y)$ -neutral.

Claim 1 Neither \mathbf{T}^* nor $\mathbf{T}^\#$ dictates that truth behaves like a classical concept in M . Proof: Say that a classical hypothesis h is *interesting* iff $h(\exists x(Gx \ \& \ \mathbf{T}x)) = h(\mathbf{T}b) = h(\neg\mathbf{T}b) = \mathbf{t}$ and $h(B) = \mathbf{f}$ for every $B \in Z$. Then, for any interesting classical hypothesis h , if $k \geq 2$ then $\tau_M^k(h)(\mathbf{T}^{k-1} A) = \tau_M^k(h)(\neg^{k-1} \mathbf{T}b) = \tau_M^k(h)(\exists x(Gx \ \& \ \mathbf{T}x)) = \mathbf{t}$; and $\tau_M^k(h)(\neg^k \mathbf{T}b) = \tau_M^k(h)(\mathbf{T}^n A) = \mathbf{f}$, where $n \neq k - 1$. So we can construct a τ_M -sequence \mathcal{S} for M such that \mathcal{S}_λ is interesting for every limit ordinal λ and such that the value of $\mathbf{T}b$ never stabilizes. In fact we can assure that $\mathbf{T}b$ is not even nearly stable.

Claim 2 For every $B \in Y$, either $B \in \mathbf{V}_M^c$ or $\neg B \in \mathbf{V}_M^c$. Proof: It suffices to show that every sentence in Y is stably \mathbf{f} in every maximally consistent τ_M -sequence \mathcal{S} . So suppose that \mathcal{S} is a maximally consistent τ_M -sequence. Then $\mathcal{S}_{n+1}(A) = \mathbf{f}$, for each $n \geq 0$, by the strong consistency of \mathcal{S}_n . So $\mathcal{S}_{k+1}(\mathbf{T}^n A) = \mathbf{f}$, for $k \geq 0$ and $n \leq k$. So $\mathcal{S}_\omega(\mathbf{T}^n A) = \mathbf{f}$ for every $n \geq 0$. So $\mathcal{S}_{\omega+1}(\exists x(Gx \ \& \ \mathbf{T}x) \ \& \ \neg\mathbf{T}b) = \mathcal{S}_{\omega+1}(\exists x(Gx \ \& \ \mathbf{T}x)) = \mathcal{S}_{\omega+1}(\mathbf{T}^n A) = \mathbf{f}$, for every $n \geq 0$. So $\mathcal{S}_{\omega+2}(\mathbf{T}b) = \mathcal{S}_{\omega+2}(\exists x(Gx \ \& \ \mathbf{T}x) \ \& \ \neg\mathbf{T}b) = \mathcal{S}_{\omega+2}(\exists x(Gx \ \& \ \mathbf{T}x)) = \mathcal{S}_{\omega+2}(\mathbf{T}^n A) = \mathbf{f}$, for every $n \geq 0$. So for every $\alpha \geq \omega + 2$ and every $n \geq 0$, we have $\mathcal{S}_\alpha(\mathbf{T}b) = \mathcal{S}_\alpha(\exists x(Gx \ \& \ \mathbf{T}x) \ \& \ \neg\mathbf{T}b) = \mathcal{S}_\alpha(\exists x(Gx \ \& \ \mathbf{T}x)) = \mathcal{S}_\alpha(\mathbf{T}^n A) = \mathbf{f}$. So every sentence in Y is stably \mathbf{f} in \mathcal{S} .

Claim 3 $lfp(\sigma 1_M)$ is classical. Proof: It suffices, given Theorem 4.24, to prove that $lfp(\sigma 1_M)(B) = \mathbf{f}$ for every $B \in Y$. Let \mathcal{S} be the $\sigma 1_M$ -sequence that iteratively builds $lfp(\sigma 1)$ from the null hypothesis: i.e., $\mathcal{S}_0(B) = \mathbf{n}$ for every $B \in S$. Note that $\mathcal{S}_{k+1}(A) = \mathbf{f}$, for $k \geq 0$. The reason is that in calculating $\mathcal{S}_{k+1}(A)$, we consider only weakly consistent classical $h \geq \mathcal{S}_k$. So $\mathcal{S}_{k+1}(\mathbf{T}^n A) = \mathbf{f}$, for $k \geq 0$ and $n \leq k$. So $\mathcal{S}_\omega(\mathbf{T}^n A) = \mathbf{f}$, for every $n \geq 0$. So, as in the proof of Claim 2, for every $\alpha \geq \omega + 2$ and every $n \geq 0$, we have $\mathcal{S}_\alpha(\mathbf{T}b) = \mathcal{S}_\alpha(\exists x(Gx \ \& \ \mathbf{T}x) \ \& \ \neg\mathbf{T}b) = \mathcal{S}_\alpha(\exists x(Gx \ \& \ \mathbf{T}x)) = \mathcal{S}_\alpha(\mathbf{T}^n A) = \mathbf{f}$. So $lfp(\sigma 1_M)(B) = \mathbf{f}$ for every $B \in Y$, as desired.

Claim 4 $lfp(\sigma 2_M)$ is classical. Proof: Note that $lfp(\sigma 1_M) \leq lfp(\sigma 2_M)$. (See the proof of Theorem 4.2.) So $lfp(\sigma 2_M)$ is classical, given Claim 3.

Example 5.10 (Gupta) This example will show that $\mathbf{T}^{lfp,\sigma} \not\leq_3 \mathbf{T}^{lfp,\sigma}$. Consider a ground language \mathcal{L} with a one-place predicate G and no other nonlogical vocabulary besides quote names. Let $D = S \cup \mathbb{N}$. For each $Y \subseteq S$, let $Y^* =$

$\{A : \neg A \in Y\}$. For each $Y \subseteq D$, we will use the notation $[Y]$ for the ground model $\langle D, I_Y \rangle$ where,

$$I_Y(G)(d) = \begin{cases} \mathbf{t}, & \text{if } d \in Y \\ \mathbf{f}, & \text{if } d \notin Y. \end{cases}$$

For nonintersecting $U, V \subseteq D$, we will use the notation (U, V) for the hypothesis h such that

$$h(d) = \begin{cases} \mathbf{t}, & \text{if } d \in U \\ \mathbf{f}, & \text{if } d \in V \\ \mathbf{n}, & \text{otherwise.} \end{cases}$$

We will define a jump operator, ϕ , not on hypotheses, but rather on subsets of S . For each $Y \subseteq S$,

$$\phi(Y) =_{\text{df}} \{A \in S : \text{Val}_{[Y \cup \mathbb{N}] + (Y, Y^* \cup \mathbb{N}), \sigma}(A) = \mathbf{t}\}.$$

Though ϕ is not in any sense monotone, it will come in handy, as we shall see. Let $Y_0 =_{\text{df}} \emptyset$. Let $Y_{n+1} =_{\text{df}} \phi(Y_n)$. And let $Y_\omega =_{\text{df}} \{A : (\exists n \geq 0)(\forall m \geq n)(A \in Y_m)\} = \bigcup_{n \geq 0} \bigcap_{m \geq n} Y_m$.

Below we will prove that the hypothesis $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is not classical, and is the least fixed point of $\sigma_{[Y_\omega \cup \mathbb{N}]}$. Given this, note that the ground model $[Y_\omega \cup \mathbb{N}]$ is $(S - Y_\omega)$ -neutral and that $\text{lfp}(\sigma_{[Y_\omega \cup \mathbb{N}]})(B) = \mathbf{t}$ for every $B \in Y_\omega$. So $\mathbf{T}^{\text{lfp}, \sigma} \not\leq_3 \mathbf{T}^{\text{lfp}, \sigma}$, as desired.

Our argument that $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is not classical, and is the least fixed point of $\sigma_{[Y_\omega \cup \mathbb{N}]}$, proceeds through numbered claims.

Claim 1 $\forall x(\mathbf{T}x \supset Gx) \notin Y_n$ and $\neg \forall x(\mathbf{T}x \supset Gx) \notin Y_n$, for each $n \geq 0$. Proof: by induction on n . The base case is vacuously true. For the inductive step, assume that $\forall x(\mathbf{T}x \supset Gx) \notin Y_n$ and $\neg \forall x(\mathbf{T}x \supset Gx) \notin Y_n$. To show that $\forall x(\mathbf{T}x \supset Gx) \notin Y_{n+1}$ and $\neg \forall x(\mathbf{T}x \supset Gx) \notin Y_{n+1}$, it suffices to show that $\text{Val}_{[Y_n \cup \mathbb{N}] + (Y_n, Y_n^* \cup \mathbb{N}), \sigma}(\forall x(\mathbf{T}x \supset Gx)) = \mathbf{n}$. Consider the classical hypotheses $h = (Y_n, D - Y_n)$ and $h' = (Y_n \cup \{\forall x(\mathbf{T}x \supset Gx)\}, D - Y_n - \{\forall x(\mathbf{T}x \supset Gx)\})$. By the inductive hypothesis, we have $(Y_n, Y_n^* \cup \mathbb{N}) \leq h, h'$. Furthermore, $\text{Val}_{[Y_n \cup \mathbb{N}] + h, \tau}(\forall x(\mathbf{T}x \supset Gx)) = \mathbf{t}$ and $\text{Val}_{[Y_n \cup \mathbb{N}] + h', \tau}(\forall x(\mathbf{T}x \supset Gx)) = \mathbf{f}$. So $\text{Val}_{[Y_n \cup \mathbb{N}] + (Y_n, Y_n^* \cup \mathbb{N}), \sigma}(\forall x(\mathbf{T}x \supset Gx)) = \mathbf{n}$, as desired.

Claim 2 $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is not classical. Proof: Given Claim 1, $\forall x(\mathbf{T}x \supset Gx) \notin Y_\omega$ and $\forall x(\mathbf{T}x \supset Gx) \notin Y_\omega^*$.

Before we state Claim 3, we define $X_n = S - (Y_n \cup Y_n^*)$ and $X_\omega = S - (Y_\omega \cup Y_\omega^*)$.

Claim 3 For each $n \geq 1$ and $m \geq 0$, there is some sentence of degree m in Y_n and some sentence of degree m in X_n . Proof: Note that $(\mathbf{T}^m A \vee \neg \mathbf{T}^m A) \in Y_n$ and $(\mathbf{T}^m A \vee \neg \mathbf{T}^m A) \& \forall x(\mathbf{T}x \supset Gx) \in X_n$, for any sentence A .

Before we state Claim 4, we introduce some notation. For $U, V \subseteq S$, say that $U =_n V$ iff for every sentence A of degree $< n$, $A \in U$ iff $A \in V$.

Claim 4 For every $n \geq 0$, and every $m \geq n + 1$, $Y_{n+1} =_n Y_m$. Proof: By induction on n . The base case is vacuously true. For the inductive step, assume that $Y_{n+1} =_n Y_m$. We want to show that $Y_{n+2} =_{n+1} Y_{m+1}$. It suffices to construct an N -restricted isomorphism Ψ from $[Y_{n+1} \cup \mathbb{N}]$ to $[Y_m \cup \mathbb{N}]$, where $N = \{ 'A' : \text{deg}(A) < n \}$. Define seven subsets of S as follows:

$$\begin{aligned}
 U &=_{\text{df}} \{ A \in S : \text{deg}(A) < n \} \\
 V &=_{\text{df}} \{ A \in S : \text{deg}(A) \geq n \ \& \ A \in Y_{n+1} \} \\
 W &=_{\text{df}} \{ A \in S : \text{deg}(A) \geq n \ \& \ A \in Y_{n+1}^* \} \\
 Z &=_{\text{df}} \{ A \in S : \text{deg}(A) \geq n \ \& \ A \in X_{n+1} \} \\
 V' &=_{\text{df}} \{ A \in S : \text{deg}(A) \geq n \ \& \ A \in Y_m \} \\
 W' &=_{\text{df}} \{ A \in S : \text{deg}(A) \geq n \ \& \ A \in Y_m^* \} \\
 Z' &=_{\text{df}} \{ A \in S : \text{deg}(A) \geq n \ \& \ A \in X_m \}
 \end{aligned}$$

Note that each of $V, W, Z, V', W',$ and Z' is countably infinite, by Claim 3. Also note that

$$\begin{aligned}
 S - U &= V \dot{\cup} W \dot{\cup} Z \\
 S - U &= V' \dot{\cup} W' \dot{\cup} Z' \\
 Y_{n+1} \cap U &= Y_m \cap U \\
 Y_{n+1}^* \cap U &= Y_m^* \cap U
 \end{aligned}$$

Define Ψ by patching together the identity function on $U \cup \mathbb{N}$, and bijections from V onto V' , from W onto W' , and from Z onto Z' .

Claim 5 $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is a fixed point of $\sigma_{[Y_\omega \cup \mathbb{N}]}$. Proof: It suffices to show that Y_ω is a fixed point of ϕ . For this, it suffices to show that $\phi(Y_\omega) =_{n+1} Y_\omega$, for every $n \geq 0$. Given Claim 4, $Y_\omega =_{n+1} Y_{n+2}$, for every $n \geq 0$. So it suffices to show that $\phi(Y_\omega) =_{n+1} Y_{n+2}$, for every $n \geq 0$. Choose any $n \geq 0$. Note that $Y_\omega =_n Y_{n+1}$, by Claim 4. To show that $\phi(Y_\omega) =_{n+1} Y_{n+2}$, it suffices to construct an N -restricted isomorphism from $[Y_\omega \cup \mathbb{N}]$ to $[Y_{n+1} \cup \mathbb{N}]$, where $N = \{ 'A' : \text{deg}(A) < n \}$. The construction follows the lines of the construction in the proof of Claim 4.

Claim 6 $(Y_\omega, Y_\omega^* \cup \mathbb{N}) = \text{lfp}(\sigma_{[Y_\omega \cup \mathbb{N}]})$. Proof: Let $(Z, Z^* \cup \mathbb{N}) = \text{lfp}(\sigma_{[Y_\omega \cup \mathbb{N}]})$. For Claim 6, it suffices to show by induction on n that $Y_\omega =_n Z$, for each $n \geq 0$. The base case is vacuously true. For the inductive step, suppose that $Y_\omega =_n Z$. We want to show that $Y_\omega =_{n+1} Z$. Note, incidentally, that $Y_\omega^* =_n Z^*$.

$Z \subseteq Y_\omega$, since $(Z, Z^* \cup \mathbb{N}) = \text{Ifp}(\sigma_{[Y_\omega \cup \mathbb{N}]}) \leq (Y_\omega, Y_\omega^* \cup \mathbb{N})$. So it suffices to show that for every sentence A of degree $< n + 1$, if $A \notin Z$ then $A \notin Y_\omega$. So suppose that $\text{deg}(A) < n + 1$ and $A \notin Z$. Then there is some classical hypothesis $(X, D - X) \geq (Z, Z^* \cup \mathbb{N})$ such that A is false in the classical model $[Y_\omega \cup \mathbb{N}] + (X, D - X)$. To show that $A \notin Y_\omega$, we will construct a classical hypothesis $(W, D - W) \geq (Y_\omega, Y_\omega^* \cup \mathbb{N})$ such that A is false in the classical model $[Y_\omega \cup \mathbb{N}] + (W, D - W)$. First we construct $(W, D - W)$, and then we construct an N -restricted isomorphism Ψ from $[Y_\omega \cup \mathbb{N}] + (X, D - X)$ to $[Y_\omega \cup \mathbb{N}] + (W, D - W)$, where $N = \{ 'B' : \text{deg}(B) < n \}$.

Define seven disjoint subsets of S , as follows:

$$\begin{aligned}
 U &=_{\text{df}} \{ A \in S : \text{deg}(A) < n \} \\
 V_1 &=_{\text{df}} (X \cap Y_\omega) - U \\
 V_2 &=_{\text{df}} (X \cap Y_\omega^*) - U \\
 V_3 &=_{\text{df}} X - (Y_\omega \cup Y_\omega^* \cup U) \\
 V'_1 &=_{\text{df}} ((S - X) \cap Y_\omega) - U \\
 V'_2 &=_{\text{df}} ((S - X) \cap Y_\omega^*) - U \\
 V'_3 &=_{\text{df}} (S - X) - (Y_\omega \cup Y_\omega^* \cup U)
 \end{aligned}$$

Note the following:

$$\begin{aligned}
 X &= V_1 \dot{\cup} V_2 \dot{\cup} V_3 \dot{\cup} (U \cap X) \\
 S - X &= V'_1 \dot{\cup} V'_2 \dot{\cup} V'_3 \dot{\cup} (U \cap (S - X)) \\
 V_3 \dot{\cup} V'_3 &= S - (Y_\omega \cup Y_\omega^* \cup U) \\
 Y_\omega - U &= V_1 \dot{\cup} V'_1 \\
 Y_\omega^* - U &= V_2 \dot{\cup} V'_2 \\
 Y_\omega \cap U &\subseteq X \cap U, \text{ since } Y_\omega =_n Z \\
 Y_\omega^* \cap U &\subseteq (S - X) \cap U, \text{ since } Y_\omega^* =_n Z^* \\
 Z - U &\subseteq V_1 \\
 Z^* - U &\subseteq V'_1
 \end{aligned}$$

Note also that each of the following sets contains sentences of arbitrarily large degree: Z, Z^* , and $S - (Y_\omega \cup Y_\omega^*)$. So each of the following sets is countably infinite: V_1, V'_2 and $V_3 \dot{\cup} V'_3$.

Choose $P \subseteq V_3$ and $Q \subseteq V'_3$, so that $P \dot{\cup} Q$ has the same cardinality as $V_2 \dot{\cup} V_3$. And let $P_1 = V_3 - P$ and $Q_1 = V'_3 - Q$. Finally, let J be a set of even numbers of the same cardinality as V'_1 . And let $K = \mathbb{N} - J$. K is countably infinite.

Let $W = (X \cap U) \dot{\cup} V_1 \dot{\cup} V'_1 \dot{\cup} P \dot{\cup} Q$. Then

$$\begin{aligned}
 S - W &= ((S - X) \cap U) \dot{\cup} V_2 \dot{\cup} V'_2 \dot{\cup} P_1 \dot{\cup} Q_1. \\
 \text{So, } Y_\omega &= (Y_\omega \cap U) \dot{\cup} V_1 \dot{\cup} V'_1 \subseteq W, \\
 \text{and } Y_\omega^* &= ((S - X) \cap U) \dot{\cup} V_2 \dot{\cup} V'_2 \subseteq S - W. \\
 \text{So, } (W, D - W) &\geq (Y_\omega, Y_\omega^* \cup \mathbb{N}).
 \end{aligned}$$

Construct an N -restricted isomorphism from $M_X = [Y_\omega \cup \mathbb{N}] + (X, D - X)$ to $M_W = [Y_\omega \cup \mathbb{N}] + (W, D - W)$ by patching together

- the identity function on U ,
- a bijection from V_1 onto $Y_\omega - U = V_1 \dot{\cup} V'_1$,
- a bijection from $V_2 \dot{\cup} V_3$ onto $P \dot{\cup} Q$,
- a bijection from $V'_2 \dot{\cup} P_1 \dot{\cup} Q_1$ onto $V_2 \dot{\cup} V'_2 \dot{\cup} P_1 \dot{\cup} Q_1$,
- a bijection from V'_1 onto J , and
- a bijection from $\mathbb{N} = J \dot{\cup} K$ onto K .

To see that Ψ is an N -restricted isomorphism from M_X to M_W , first note that Ψ maps the extension of G in M_X one-one onto the extension of G in M_W . The reason is that $Y_\omega \cup \mathbb{N} = (U \cap Y_\omega) \dot{\cup} V_1 \dot{\cup} V'_1 \dot{\cup} J \dot{\cup} K$ and Ψ maps V_1 one-one onto $V_1 \dot{\cup} V'_1$, and V'_1 one-one onto J , and $J \dot{\cup} K$ one-one onto K . Also, Ψ maps $X = (U \cap X) \dot{\cup} V_1 \dot{\cup} V_2 \dot{\cup} V_3$ one-one onto $W = (U \cap X) \dot{\cup} V_1 \dot{\cup} V'_1 \dot{\cup} P \dot{\cup} Q$, since Ψ maps V_1 one-one onto $V_1 \dot{\cup} V'_1$, and $V_2 \dot{\cup} V_3$ one-one onto $P \dot{\cup} Q$. So Ψ maps the extension of T in M_X one-one onto the extension of T in M_W . Finally, note that for every name ‘ B ’ in N , Ψ maps the denotation of ‘ B ’ in M_X to the denotation of ‘ B ’ in M_W , since, $\Psi(B) = B$ for every $B \in U$.

Example 5.11 (Gupta) Here we modify Example 5.10 to get a proof that $\mathbf{T}^{lfp,\sigma^1} \not\leq_3 \mathbf{T}^{lfp,\sigma^1}$. As we shall see, our modified example will also show that $\mathbf{T}^{lfp,\sigma^2} \not\leq_2 \mathbf{T}^{lfp,\sigma^1}$. The current example starts like Example 5.10, except that the jump operator ϕ must now be defined for $Y \subseteq S$ as follows:

$$\phi(Y) =_{\text{df}} \{A \in S : \text{Val}_{[Y \cup \mathbb{N}] + (Y, Y^* \cup \mathbb{N}), \sigma^1}(A) = \mathbf{t}\}.$$

As in Example 5.10, $Y_0 =_{\text{df}} \emptyset$, $Y_{n+1} =_{\text{df}} \phi(Y_n)$, and $Y_\omega =_{\text{df}} \bigcup_{n \geq 0} \bigcap_{m \geq n} Y_m$.

Below, we state six claims analogous to the claims made in Example 5.10. Except where indicated, their proofs are exactly as in Example 5.10, except that occurrences of “ σ ” must be replaced by “ σ^1 ”. Before we state our new Claims 1-6, we state a Claim 0.

Claim 0 $Y_n \cup \{\forall x(Tx \supset Gx)\}$ is consistent for every $n \geq 0$. Proof: by induction on $n \geq 0$. The base case is trivially true. For the inductive step, suppose that $Y_n \cup \{\forall x(Tx \supset Gx)\}$ is consistent. Note that every sentence in $Y_{n+1} \cup \{\forall x(Tx \supset Gx)\}$ is true in the classical model $[Y_n \cup \mathbb{N}] + (Y_n, D - Y_n)$. So $Y_{n+1} \cup \{\forall x(Tx \supset Gx)\}$ is consistent.

Claim 1 $\forall x(Tx \supset Gx) \notin Y_n$ and $\neg \forall x(Tx \supset Gx) \notin Y_n$, for each $n \geq 0$. The proof proceeds exactly as in Example 5.4, except that after introducing $h = (Y_n, D - Y_n)$ and $h' = (Y_n \cup \{\forall x(Tx \supset Gx)\}, D - Y_n - \{\forall x(Tx \supset Gx)\})$, we have to verify that h and h' are not only classical but also weakly consistent. But this follows from Claim 0.

Claim 2 $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is not classical.

Claim 3 For each $n \geq 1$ and $m \geq 0$, there is some sentence of degree m in Y_n and some sentence of degree m in X_n .

Claim 4 For every $n \geq 0$, and every $m \geq n + 1$, $Y_{n+1} =_n Y_m$.

Claim 5 $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is a fixed point of $\sigma 1_{[Y_\omega \cup \mathbb{N}]}$.

Proof of Claims 2–5 Exactly as in Example 5.10. □

Claim 6 $(Y_\omega, Y_\omega^* \cup \mathbb{N}) = \text{lfp}(\sigma 1_{[Y_\omega \cup \mathbb{N}]})$. Proof: we must modify the construction in the proof of Example 5.10, Claim 6, as follows. First, replace the fourth and fifth sentences of the second paragraph of the proof of Claim 6 in Example 5.10 with the following:

Then there is some *weakly consistent* classical hypothesis $(X, D - X) \geq (Z, Z^* \cup \mathbb{N})$ such that A is false in the classical model $[Y_\omega \cup \mathbb{N}] + (X, D - X)$. To show that $A \notin Y_\omega$, we will construct a *weakly consistent* classical hypothesis $(W, D - W) \geq (Y_\omega, Y_\omega^* \cup \mathbb{N})$ such that A is false in the classical model $[Y_\omega \cup \mathbb{N}] + (W, D - W)$.

Then, up until the choice of $P \subseteq V_3$ and $Q \subseteq V'_3$, the construction proceeds exactly as in Example 5.10. But before we choose P and Q , we must first prove that $(X \cap U) \cup Y_\omega = (X \cap U) \dot{\cup} V_1 \dot{\cup} V'_1$ is consistent. Suppose not. Then, by compactness, $Y_\omega \cup \{B_1, \dots, B_k\}$ is inconsistent for some $B_1, \dots, B_k \in (X \cap U)$. So Y_ω logically implies $B =_{\text{df}} \neg(B_1 \ \& \ \dots \ \& \ B_k)$. So $B \in Y$. Note that $\text{deg}(B) < n$, since each $B_i \in U$. So $B \in Z$, since $Y_\omega =_n Z$. But $Z \subseteq X$ and $B_1, \dots, B_k \in X$. So X is inconsistent. So $(X, D - X)$ is not weakly consistent, a reductio. So $(X \cap U) \cup Y_\omega$ is consistent, as desired.

Now we will choose $P \subseteq V_3$ and $Q \subseteq V'_3$, but more carefully than in Example 5.10. Note that $V_3 \dot{\cup} V'_3$ contains countably infinitely many sentences and is closed under negation. Also, $(X \cap U) \cup Y_\omega$ is consistent. So there are closedly infinitely many sentences in $V_3 \dot{\cup} V'_3$ that are consistent with $(X \cap U) \cup Y_\omega$. So we can choose $P \subseteq V_3$ and $Q \subseteq V'_3$ so that $(X \cap U) \dot{\cup} V_1 \dot{\cup} V'_1 \dot{\cup} P \dot{\cup} Q = (X \cap U) \dot{\cup} Y_\omega \dot{\cup} P \dot{\cup} Q$ is consistent and so that $P \dot{\cup} Q$ has the same cardinality as $V_2 \dot{\cup} V_3$.

Let $W = (X \cap U) \dot{\cup} V_1 \dot{\cup} V'_1 \dot{\cup} P \dot{\cup} Q$, as in Example 5.10. W is consistent. So the hypothesis $(W, D - W)$ is weakly consistent. The construction of the

restricted isomorphism Ψ goes through as in Example 5.10. So A is false in the classical model $[Y_\omega \cup \mathbb{N}] + (W, D - W)$, as desired.

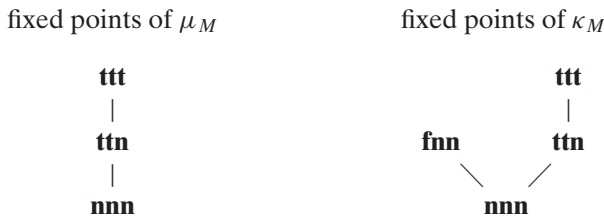
So $(Y_\omega, Y_\omega^* \cup \mathbb{N})$ is not classical, and is the least fixed point of $\sigma 1_{[Y_\omega \cup \mathbb{N}]}$. But note that the ground model $[Y_\omega \cup \mathbb{N}]$ is $(S - Y_\omega)$ -neutral, and $lfp(\sigma 1_{[Y_\omega \cup \mathbb{N}]})(B) = \mathbf{t}$ for every $B \in Y_\omega$. So $\mathbf{T}^{lfp,\sigma^1} \not\leq_3 \mathbf{T}^{lfp,\sigma^1}$, as desired.

We furthermore claim that $lfp(\sigma 2_{[Y_\omega \cup \mathbb{N}]})$ is classical. Firstly, $lfp(\sigma 1_{[Y_\omega \cup \mathbb{N}]}) \leq lfp(\sigma 2_{[Y_\omega \cup \mathbb{N}]})$. So the ground model $[Y_\omega \cup \mathbb{N}]$ is $(S - Y_\omega)$ -neutral, and $lfp(\sigma 2_{[Y_\omega \cup \mathbb{N}]})(B) = \mathbf{t}$ for every $B \in Y_\omega$. So $lfp(\sigma 2_{[Y_\omega \cup \mathbb{N}]})$ is classical, since $\mathbf{T}^{lfp,\sigma^2} \leq_3 \mathbf{T}^{lfp,\sigma^2}$ by Theorem 4.21 (2), proved above. So $\mathbf{T}^{lfp,\sigma^2} \leq_2 \mathbf{T}^{lfp,\sigma^1}$.

Example 5.12 This example will show that $\mathbf{T}^{gifp,\mu} \not\leq_2 \mathbf{T}^{gifp,\kappa}$. Consider a ground language \mathcal{L} with exactly two nonquote names, b and c , and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with $I(b) = B = \mathbf{T}b \ \& \ \mathbf{T}c$ and $I(c) = C = \mathbf{T}b \vee \neg \mathbf{T}c$. The facts in the following table can be easily established by calculating:

If $\langle h(B), h(C) \rangle$	=	tt	tf	tn	ft	ff	fn	nt	nf	nn
then $\langle \mu_M(h)(B), \mu_M(h)(C) \rangle$	=	tt	ft	nn	ff	ft	nn	nn	nn	nn
and $\langle \kappa_M(h)(B), \kappa_M(h)(C) \rangle$	=	tt	ft	nt	ff	ft	fn	nn	ft	nn

Given this table, we can argue as in Gupta and Belnap’s Transfer Theorem ([3], Theorem 2D.4) to the following conclusion: μ_M has three fixed points, which are completely determined by the ordered triple $\langle h(B), h(C), h(\forall x(\mathbf{T}x \vee \neg \mathbf{T}x)) \rangle$; and κ_M has three fixed points, which are completely determined by the ordered triple $\langle h(B), h(C), h(\forall x(\mathbf{T}x \vee \neg \mathbf{T}x)) \rangle$. Furthermore τ_M has exactly one fixed point, in which both B and C are \mathbf{t} . Finally, that unique fixed point of τ_M is also a fixed point of μ_M and of κ_M . The fixed points of μ_M and κ_M line up as follows:



Thus $gifp(\mu_M)$ is classical but $gifp(\kappa_M)$ is not.

Example 5.13 This example will show that $\mathbf{T}^{gifp,\kappa} \not\leq_2 \mathbf{T}^{gifp,\sigma}$. Consider a ground language \mathcal{L} with exactly two nonquote names, b and c , and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with

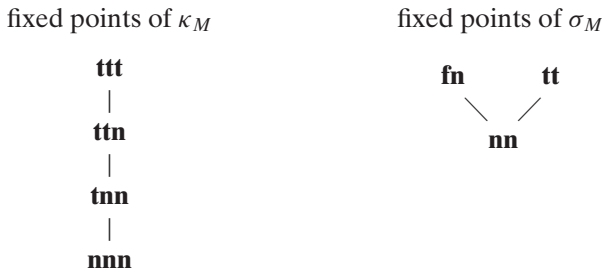
$$I(b) = B = \mathbf{T}b \vee (\mathbf{T}c \ \& \ \neg \mathbf{T}c), \text{ and}$$

$$I(c) = C = (\mathbf{T}b \ \& \ (\mathbf{T}c \vee \neg \mathbf{T}c)) \vee (\neg \mathbf{T}b \ \& \ \neg \mathbf{T}c).$$

The facts in the following table can be easily established by calculating:

If $\langle h(B), h(C) \rangle$	=	tt	tf	tn	ft	ff	fn	nt	nf	nn
then $\langle \kappa_M(h)(B), \kappa_M(h)(C) \rangle$	=	tt	tt	tn	ff	ft	nn	nn	nn	nn
and $\langle \sigma_M(h)(B), \sigma_M(h)(C) \rangle$	=	tt	tt	tt	ff	ft	fn	nn	nt	nn

Given this table, we can argue as in Gupta and Belnap’s Transfer Theorem ([3], Theorem 2D.4) to the following conclusion: κ_M has three fixed points, which are completely determined by the ordered triple $\langle h(B), h(C), h(\forall x(Tx \vee \neg Tx)) \rangle$; and σ_M has three fixed points, which are completely determined by the ordered pair $\langle h(B), h(C) \rangle$. (The reason we only need look at these *pairs* of truth values is that the proviso in the Transfer Theorem can be dropped for σ .) Furthermore τ_M has exactly one fixed point, in which both B and C are **t**. Finally, that unique fixed point of τ_M is also a fixed point of κ_M and of σ_M . The fixed points of κ_M and σ_M line up as follows:



Thus $gifp(\kappa_M)$ is classical but $gifp(\sigma_M)$ is not.

Example 5.14 This example will show that $\mathbf{T}^{gifp,\sigma} \not\leq_2 \mathbf{T}^{gifp,\sigma^1}$. Consider a ground language \mathcal{L} with exactly four nonquote names, b, c, d and e , and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with

$$\begin{aligned}
 I(b) &= B = Tb \vee (Td \ \& \ Te) \\
 I(c) &= C = Tb \vee \neg Tc \\
 I(d) &= D = Tc \\
 I(e) &= E = \neg Tc
 \end{aligned}$$

The facts in the following table can be established by calculating. The asterisks are classical wildcards, either **t** or **f** and the question marks can vary with the wildcards:

If $\langle h(B), h(C), h(D), h(E) \rangle$	=	tt**	ft**	*f**
then $\langle \tau_M(h)(B), \tau_M(h)(C), \tau_M(h)(D), \tau_M(h)(E) \rangle$	=	tttf	?ftf	?ftf

From Gupta and Belnap’s Transfer Theorem, we conclude that τ_M has a unique fixed point, say h_0 , where $h_0(B) = h_0(C) = h_0(D) = \mathbf{t}$ and $h_0(E) = \mathbf{f}$. Since h_0 is a fixed point of τ_M , it is also a fixed point of σ_M and of $\sigma 1_M$.

Furthermore, by an argument similar to that given for the Transfer Theorem, we can conclude that the fixed points of σ_M are completely determined by the values $\langle h(B), h(C), h(D), h(E) \rangle$. (The reason we only need look at these quartuples of truth values is that the proviso in the Transfer Theorem can be dropped for σ .) Thus, h_0 is the only classical fixed point of σ_M , and the only fixed point of σ_M in which B, C and D are \mathbf{t} and E is \mathbf{f} . In fact, h_0 is the only fixed point of σ_M in which B and C are \mathbf{t} , since in any such fixed point D is \mathbf{t} and E is \mathbf{f} .

Claim 1 σ_M has no fixed point h where $h(C) = \mathbf{f}$. Proof: Suppose that h is a fixed point of σ_M with $h(C) = \mathbf{f}$. Then, since h is a fixed point of σ_M , $h(\mathbf{T}c) = h(\mathbf{T}'c) = h(C) = \mathbf{f}$. So $h(C) = h(\mathbf{T}b \vee \neg\mathbf{T}c) = \text{Val}_{M+h,\sigma}(\mathbf{T}b \vee \neg\mathbf{T}c) = \mathbf{t}$, a contradiction.

Claim 2 σ_M has no fixed point h where $h(B) = \mathbf{f}$. Proof: Suppose that h is a fixed point of σ_M with $h(B) = \mathbf{f}$. By Claim 1, $h(C) = \mathbf{t}$ or \mathbf{n} . If $h(C) = \mathbf{t}$ then, since h is a fixed point of σ_M , $h(\mathbf{T}c) = \mathbf{t}$, so that $h(C) = h(\mathbf{T}b \vee \neg\mathbf{T}c) = \text{Val}_{M+h,\sigma}(\mathbf{T}b \vee \neg\mathbf{T}c) = \mathbf{f}$, a contradiction. So $h(C) = \mathbf{n}$. So $h(D) = h(\mathbf{T}c) = \mathbf{n} = h(\neg\mathbf{T}c) = h(E)$. Let $h' \geq h$ be a classical hypothesis with $h'(\mathbf{T}c) = h'(\neg\mathbf{T}c) = \mathbf{t}$, and let $h'' \geq h$ be a classical hypothesis with $h''(\mathbf{T}c) = h''(\neg\mathbf{T}c) = \mathbf{f}$. Then $\text{Val}_{M+h',\tau}(\mathbf{T}d) = \text{Val}_{M+h',\tau}(\mathbf{T}e) = \mathbf{t}$ and $\text{Val}_{M+h'',\tau}(\mathbf{T}d) = \text{Val}_{M+h'',\tau}(\mathbf{T}e) = \mathbf{f}$. So $\tau_M(h')(B) = \text{Val}_{M+h',\tau}(\mathbf{T}b \vee (\mathbf{T}d \ \& \ \mathbf{T}e)) = \mathbf{t}$ and $\tau_M(h'')(B) = \text{Val}_{M+h'',\tau}(\mathbf{T}b \vee (\mathbf{T}d \ \& \ \mathbf{T}e)) = \mathbf{f}$. So $\sigma_M(h)(B) = \mathbf{n}$. This contradicts h ’s being a fixed point of σ_M with $h(B) = \mathbf{f}$.

Given Claims 1 and 2, σ_M has no fixed points that are incompatible with h_0 . So h_0 is a σ_M -intrinsic fixed point. So $h_0 = \text{gifp}(\sigma_M)$, since h_0 is classical.

As for $\sigma 1$, let g be the (weakly consistent) hypothesis with $g(B) = \mathbf{f}$, and $g(A) = \mathbf{n}$ for every $A \in S - \{B\}$. Note that $\sigma 1_M(g)(B) = \mathbf{f}$. So $g \leq \sigma 1_M(g)$. By the monotony of $\sigma 1_M$, there is a unique $\sigma 1_M$ -sequence \mathcal{S} with $\mathcal{S}_0 = g$. Furthermore, \mathcal{S} is (nonstrictly) increasing, and culminates in a fixed point of $\sigma 1_M$, say h_1 . Note that $h_1(B) = \mathbf{f}$. But h_0 is also a fixed point of $\sigma 1_M$, and $h_0(B) = \mathbf{t}$. So $\text{gifp}(\sigma 1_M)$ is not classical.

Example 5.15 This example will show that $\mathbf{T}^{\text{gifp},\sigma 1} \not\leq_2 \mathbf{T}^{\text{gifp},\sigma 2}$. Consider a ground language \mathcal{L} with exactly four nonquote names, b, c, d and e , and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with

$$\begin{aligned} I(b) &= B = \mathbf{T}b \vee (\neg\mathbf{T}d \ \& \ \neg\mathbf{T}e) \\ I(c) &= C = \mathbf{T}b \vee \neg\mathbf{T}c \\ I(d) &= D = \mathbf{T}c \\ I(e) &= E = \neg\mathbf{T}c \end{aligned}$$

The facts in the following table can be established by calculating. The asterisks are classical wildcards, either **t** or **f** and the question marks can vary with the wildcards:

If $\langle h(B), h(C), h(D), h(E) \rangle$	=	tt**	ft**	*f**
then $\langle \tau_M(h)(B), \tau_M(h)(C), \tau_M(h)(D), \tau_M(h)(E) \rangle$	=	tttf	?ftf	?tft

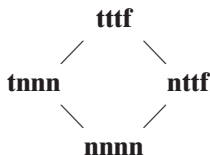
From Gupta and Belnap’s Transfer Theorem, we conclude that τ_M has a unique fixed point, say h_0 , where $h_0(B) = h_0(C) = h_0(D) = \mathbf{t}$ and $h_0(E) = \mathbf{f}$. Since h_0 is a fixed point of τ_M , it is also a fixed point of $\sigma 1_M$ and of $\sigma 2_M$.

Furthermore, by an argument similar to that given for the Transfer Theorem, we can conclude that the fixed points of σ are completely determined by the values $\langle h(B), h(C), h(D), h(E) \rangle$. Thus, h_0 is the only classical fixed point of $\sigma 1_M$, and the only fixed point of $\sigma 1_M$ in which B, C and D are **t** and E is **f**. In fact, h_0 is the only fixed point of $\sigma 1_M$ in which B and C are **t**, since in any such fixed point D is **t** and E is **f**.

Claim 1 $\sigma 1_M$ has no fixed point h where $h(C) = \mathbf{f}$. Proof: Suppose that h is a fixed point of $\sigma 1_M$ with $h(C) = \mathbf{f}$. Then, since h is a fixed point of $\sigma 1_M$, $h(Tc) = h(T^cC) = h(C) = \mathbf{f}$. So $h(C) = h(Tb \vee \neg Tc) = Val_{M+h,\sigma}(Tb \vee \neg Tc) = \mathbf{t}$, a contradiction.

Claim 2 $\sigma 1_M$ has no fixed point h where $h(B) = \mathbf{f}$. Proof: Suppose that h is a fixed point of $\sigma 1_M$ with $h(B) = \mathbf{f}$. By Claim 1, $h(C) = \mathbf{t}$ or **n**. If $h(C) = \mathbf{t}$ then, since h is a fixed point of $\sigma 1_M$, $h(Tc) = \mathbf{t}$, so that $h(C) = h(Tb \vee \neg Tc) = Val_{M+h,\sigma}(Tb \vee \neg Tc) = \mathbf{f}$, a contradiction. So $h(C) = \mathbf{n}$. So $h(T^cC) = h(Tc) = \mathbf{n} = h(\neg Tc)$, since h is a fixed point. Consider the classical hypothesis h' such that, for every $A \in S$, $h'(A) = \mathbf{t}$ iff $h(A) = \mathbf{t}$. The hypothesis h' is weakly consistent, since the set $\{A \in S : h(A) = \mathbf{t}\}$ is consistent, h being a fixed point of $\sigma 1$. Also note that $h'(Tc) = h'(\neg Tc) = h'(B) = \mathbf{f}$. So $Val_{M+h',\tau}(Td) = Val_{M+h',\tau}(T^cT^c) = \mathbf{f} = Val_{M+h',\tau}(T^c\neg Tc) = Val_{M+h',\tau}(Te)$. So $Val_{M+h',\tau}(B) = Val_{M+h',\tau}(Tb \vee (\neg Td \ \& \ \neg Te)) = \mathbf{t}$. So $\sigma 1_M(h)(B) \neq \mathbf{f} = h(B)$. But this contradicts h ’s being a fixed point of $\sigma 1$.

So for every fixed point h of $\sigma 1$, the possible values for the quartuple $\langle h(B), h(C), h(D), h(E) \rangle$ are **tttf**, **tnnn**, **nttf**, and **nnnn**. As already pointed out, each fixed point h of $\sigma 1$ is uniquely determined by $\langle h(B), h(C), h(D), h(E) \rangle$; moreover the ordering on them is isomorphic to the ordering induced on the four quartuples **tttf**, **tnnn**, **nttf**, and **nnnn**:



So h_0 is the greatest fixed point of $\sigma 1_M$. So $h_0 = gifp(\sigma 1_M)$, which is classical.

As for σ_2 , let g be the (weakly consistent) hypothesis with $g(B) = \mathbf{f}$, and $g(A) = \mathbf{n}$ for every $A \in S - \{B\}$. Note that $\sigma_{2_M}(g)(B) = \mathbf{f}$. So $g \leq \sigma_{2_M}(g)$. By the monotony of σ_{2_M} , there is a unique σ_{2_M} -sequence \mathcal{S} with $\mathcal{S}_0 = g$. Furthermore, \mathcal{S} is (nonstrictly) increasing, and culminates in a fixed point of σ_{2_M} , say h_1 . Note that $h_1(B) = \mathbf{f}$. But h_0 is also a fixed point of σ_{2_M} , and $h_0(B) = \mathbf{t}$. So $\text{gifp}(\sigma_{2_M})$ is not classical.

Example 5.16 This example will show that $\mathbf{T}^{\text{gifp}, \sigma_2} \not\leq_2 \mathbf{T}^c$. Consider a ground language \mathcal{L} with exactly two nonquote names, b and c , and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with $I(b) = B = \mathbf{T}c$ and $I(c) = C = \mathbf{T}b \ \& \ \neg \mathbf{T}c$. The facts in the following table can be easily established by calculating:

If $\langle h(B), h(C) \rangle$	=	tt	tf	tn	ft	ff	fn	nt	nf	nn
then $\langle \sigma_{2_M}(h)(B), \sigma_{2_M}(h)(C) \rangle$	=	tf	ft		tf	ff				

We have not filled in all the spaces in the table. These are not trivial: in order to calculate these values, we must know which classical $h' \geq h$ are strongly consistent. Right away we know that there are no strongly consistent hypotheses h with $h(B) = h(C) = \mathbf{t}$. So we can fill in the third column with **ft**. For our purposes, we do not need all the other columns. All we need is the following:

If $\langle h(B), h(C) \rangle$	=	tt	tf	tn	ft	ff	fn	nt	nf	nn
then $\langle \sigma_{2_M}(h)(B), \sigma_{2_M}(h)(C) \rangle$	=	tf	ft	ft	tf	ff	?f	tf	??	??

By an argument similar to that given for the Transfer Theorem, the fixed points of σ_{2_M} are completely determined by the values $\langle h(B), h(C) \rangle$; and the fixed point h_0 determined by the values **ff** is classical. Furthermore, the only other potential fixed points are determined by the values **nf** and **nn**. If such fixed points exist, they are both $\leq h_0$. So, whatever other fixed points there might be, $h_0 = \text{gifp}(\sigma_{2_M})$. So $\text{gifp}(\sigma_{2_M})$ is classical.

Now we will show that \mathbf{T}^c does not dictate that truth behaves like a classical concept in M . Choose any strongly consistent hypothesis h with $h(B) = \mathbf{t}$ and $h(C) = \mathbf{f}$. This can be done since $(B \ \& \ \neg C)$ is consistent. Note that if n is even then $\tau_M^n(h)(B) = \mathbf{t}$ and $\tau_M^n(h)(C) = \mathbf{f}$; and if n is odd then $\tau_M^n(h)(B) = \mathbf{f}$ and $\tau_M^n(h)(C) = \mathbf{t}$. So there is some maximally consistent τ_M -sequence \mathcal{S} such that neither B nor C is stable in \mathcal{S} .

Example 5.17 This example will show that (1) $\mathbf{T}^{lfp, \kappa} \not\leq_1 \mathbf{T}^{lfp, \mu}$ and (2) $\mathbf{T}^{lfp, \rho'} \not\leq_1 \mathbf{T}^{\text{gifp}, \rho}$, where ρ and ρ' are chosen with ρ strictly to the left of ρ' from the list $\mu, \kappa, \sigma, \sigma_1, \sigma_2$. Consider a ground language \mathcal{L} with exactly one nonquote name, b , and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with $I(b) = \neg \mathbf{T}b$. Let

$C = \exists x(x = x)$. Note that $h(B) = \mathbf{n}$ for any fixed point h of $\mu_M, \kappa_M, \sigma_M, \sigma 1_M$ or $\sigma 2_M$. So $lfp(\kappa_M)(B \vee C) = lfp(\sigma_M)(B \vee \neg B) = lfp(\sigma 1_M)(\neg \mathbf{T}^c B' \vee \neg \mathbf{T}^c \neg B') = lfp(\sigma 2_M)(\mathbf{T}^c B' \vee \mathbf{T}^c \neg B') = \mathbf{t}$. Meanwhile $lfp(\mu_M)(B \vee C) = gifp(\mu_M)(B \vee C) = gifp(\kappa_M)(B \vee \neg B) = gifp(\sigma_M)(\neg \mathbf{T}^c B' \vee \neg \mathbf{T}^c \neg B') = gifp(\sigma 1_M)(\mathbf{T}^c B' \vee \mathbf{T}^c \neg B') = \mathbf{n}$.

Example 5.18 This example will show that $\mathbf{T}^* \not\leq_1 \mathbf{T}^{gifp,\rho}$, for $\rho = \sigma, \sigma 1, \sigma 2$. Consider a ground language \mathcal{L} with countably many nonquote names, $b_0, b_1, \dots, b_n, \dots$; a one-place predicate G ; and no other nonlogical vocabulary besides quote names. Let M be the ground model $\langle S, I \rangle$ with

$$\begin{aligned}
 I(b_0) &= B_0 = \exists x \exists y (Gx \ \& \ Gy \ \& \ \mathbf{T}x \ \& \ \mathbf{T}y \ \& \ x \neq y) \\
 &\quad \vee \forall x (Gx \supset \neg \mathbf{T}x) \vee \mathbf{T}b_0, \\
 I(b_i) &= B_i = \forall x (Gx \supset (\mathbf{T}x \equiv x = b_i)), \text{ for } i \geq 1, \text{ and}
 \end{aligned}$$

and $I(G)(A) = \mathbf{t}$ iff $A \in Y = \{B_0, B_1, \dots\}$. Note that M is $(S - Y)$ -neutral.

Let C be the sentence $\exists x(Gx \ \& \ \mathbf{T}x)$. It will suffice to show that C is stably \mathbf{t} in every τ_M -sequence, but that $gifp(\sigma 2_M)(C) = gifp(\sigma 1_M)(C) = gifp(\sigma_M)(C) = \mathbf{n}$.

For each $n \geq 0$, define a set H_n of hypotheses, and a set \mathcal{S}_n of τ_M -sequences:

$$\begin{aligned}
 H_n &=_{\text{df}} \{h : h \text{ is a classical hypothesis,} \\
 &\quad \text{and } h(B_n) = \mathbf{t} \text{ and } h(B_m) = \mathbf{f} \text{ for } m \neq n\}, \text{ and} \\
 \mathcal{S}_n &=_{\text{df}} \{\mathcal{S} : \mathcal{S} \text{ is a } \tau_M\text{-sequence and } \mathcal{S}_1 \in H_n\}
 \end{aligned}$$

First note that, for any hypothesis h , if $h \in H_n$ then $\tau_M(h) \in H_n$ and if $h \notin \cup_n H_n$ then $\tau_M(h) \in H_0$. Thus, for any τ_M -sequence \mathcal{S} , if $\mathcal{S}_0 \in H_n$ then $\mathcal{S}_1 \in H_n$ and if $\mathcal{S}_0 \notin \cup_n H_n$ then $\mathcal{S}_1 \in H_0$. Thus, for every τ_M -sequence \mathcal{S} , there is a unique $n \geq 0$ such that $\mathcal{S} \in \mathcal{S}_n$.

Claim 1 If $\mathcal{S}, \mathcal{S}' \in \mathcal{S}_n$, then \mathcal{S} and \mathcal{S}' culminate in the same fixed point, which is itself in H_n . Choose $\mathcal{S}, \mathcal{S}' \in \mathcal{S}_n$. Note that not only $\mathcal{S}_1, \mathcal{S}'_1 \in H_n$, but also $\mathcal{S}_m, \mathcal{S}'_m \in H_n$ for every $m \geq 1$. Thus $\mathcal{S}_m =_Y \mathcal{S}'_m$ for any $m, k \geq 1$. Thus, by Corollary 5.5, $\mathcal{S}_\omega = \mathcal{S}'_\omega$ is a fixed point of τ_M . And note that $\mathcal{S}_\omega \in H_n$.

So τ_M has a unique fixed point $h_n \in H_n$, for each $n \geq 0$; and every τ_M -sequence culminates in one of the h_n . Note that $h_n(C) = \mathbf{t}$, for each $n \geq 0$. So C is stably \mathbf{t} in every τ_M -sequence.

It remains to show that $gifp(\sigma 2_M)(C) = gifp(\sigma 1_M)(C) = gifp(\sigma_M)(C) = \mathbf{n}$. We will only give the argument for $gifp(\sigma 2_M)(C)$; the other arguments are similar.

Define hypotheses h^* and h' as follows:

$$h^*(A) = \begin{cases} \mathbf{t}, & \text{if } A \in \mathbf{V}_M^* \\ \mathbf{f}, & \text{if } \neg A \in \mathbf{V}_M^* \\ \mathbf{n}, & \text{otherwise.} \end{cases}$$

$$h'(A) = \begin{cases} \mathbf{t}, & \text{if } A \in \mathbf{V}_M^* \\ \mathbf{f}, & \text{if } \neg A \in \mathbf{V}_M^* \text{ or } A \in Y \\ \mathbf{n}, & \text{otherwise.} \end{cases}$$

Then h^* is the greatest lower bound of the h_n , and is strongly consistent. Also, $h^* \leq h'$: shortly, we will see that h' is also strongly consistent. Any intrinsic fixed point of $\sigma 2_M$ must be \leq any classical fixed point of τ_M . So $\text{gifp}(\sigma 2_M) \leq h_n$, for each $n \geq 0$. So $\text{gifp}(\sigma 2_M) \leq h^*$.

Now notice that, for any $n \geq 0$, the set $\mathbf{V}_M^* \cup \{\neg B_0, \neg B_1, \dots, \neg B_n\}$ is a consistent set of sentences since $\mathbf{V}_M^* \cup \{\neg B_0, \neg B_1, \dots, \neg B_n\} \subseteq \{A \in S : h_{n+1}(A) = \mathbf{t}\}$. So, by compactness, $\mathbf{V}_M^* \cup \{\neg B_0, \neg B_1, \dots, \neg B_n, \dots\}$ is consistent. So h' is strongly consistent, as recently claimed.

Since h' is strongly consistent, there is a *classical* strongly consistent $h'' \geq h'$. Thus $h'' \geq h' \geq h^* \geq \text{gifp}(\sigma 2_M)$. Note that $h''(A) = \mathbf{f}$ for each $A \in Y$. So $\tau_M(h'')(C) = \mathbf{f}$. So $\text{gifp}(\sigma 2_M)(C) = \sigma 2_M(\text{gifp}(\sigma 2_M))(C) \neq \mathbf{t}$. Also $\text{gifp}(\sigma 2_M)(C) \neq \mathbf{f}$, since $\text{gifp}(\sigma 2_M) \leq h_0$ and $h_0(C) = \mathbf{t}$. So $\text{gifp}(\sigma 2_M)(C) = \mathbf{n}$, as desired.

6 Wrapping Up

So far we have the following results:

- Positive results proved in Section 4:
 - $\mathbf{T}^{lfp,\mu} \leq_1 \mathbf{T}^{lfp,\kappa} \leq_1 \mathbf{T}^{lfp,\sigma} \leq_1 \mathbf{T}^{lfp,\sigma 1} \leq_1 \mathbf{T}^{lfp,\sigma 2}$.
 - $\mathbf{T}^{lfp,\rho} \leq_1 \mathbf{T}^{gifp,\rho}$, for $\rho = \mu, \kappa, \sigma, \sigma 1$ or $\sigma 2$.
 - $\mathbf{T}^{lfp,\sigma} \leq_1 \mathbf{T}^* \leq_1 \mathbf{T}^\#$. $\mathbf{T}^* \leq_1 \mathbf{T}^c$. $\mathbf{T}^{lfp,\sigma 2} \leq_1 \mathbf{T}^c$.
 - $\mathbf{T}^{lfp,\mu} \equiv_2 \mathbf{T}^{lfp,\kappa} \leq_2 \mathbf{T}^{lfp,\sigma} \leq_2 \mathbf{T}^{lfp,\sigma 1} \leq_2 \mathbf{T}^{lfp,\sigma 2} \leq_2 \mathbf{T}^c$.
 - $\mathbf{T}^{lfp,\sigma} \leq_2 \mathbf{T}^* \leq_2 \mathbf{T}^\#$.
 - $\mathbf{T}^* \leq_2 \mathbf{T}^c \leq_2 \mathbf{T}^{gifp,\sigma 2} \leq_2 \mathbf{T}^{gifp,\sigma 1} \leq_2 \mathbf{T}^{gifp,\sigma} \leq_2 \mathbf{T}^{gifp,\kappa} \leq_2 \mathbf{T}^{gifp,\mu}$.
 - $\mathbf{T}^* \leq_3 \mathbf{T}^c \leq_3 \mathbf{T}^{gifp,\sigma 2} \leq_3 \mathbf{T}^{gifp,\sigma 1} \leq_3 \mathbf{T}^{gifp,\sigma} \leq_3 \mathbf{T}^{gifp,\kappa} \leq_3 \mathbf{T}^{gifp,\mu}$.
 - $\mathbf{T}^{lfp,\mu} \leq_3 \mathbf{T}^{lfp,\sigma} \leq_3 \mathbf{T}^{lfp,\sigma 1} \leq_3 \mathbf{T}^{lfp,\sigma 2} \leq_3 \mathbf{T}^c$. $\mathbf{T}^{lfp,\kappa} \leq_3 \mathbf{T}^{lfp,\sigma}$.
- Positive results proved in Section 5:
 - $\mathbf{T}^* \leq_3 \mathbf{T}^\#$. $\mathbf{T}^c \leq_3 \mathbf{T}^c$. $\mathbf{T}^{lfp,\sigma 2} \leq_3 \mathbf{T}^{lfp,\sigma 2}$.
 - $\mathbf{T}^{gifp,\rho} \leq_3 \mathbf{T}^{gifp,\rho}$, for $\rho = \mu, \kappa, \sigma, \sigma 1$ or $\sigma 2$.
- Negative results from the examples in Section 5:
 - $\mathbf{T}^\# \not\leq_3 \mathbf{T}^\#$. $\mathbf{T}^\# \not\leq_2 \mathbf{T}^*$. $\mathbf{T}^\# \not\leq_2 \mathbf{T}^{gifp,\mu}$.
 - $\mathbf{T}^{lfp,\sigma 1} \not\leq_2 \mathbf{T}^*$. $\mathbf{T}^{lfp,\sigma 1} \not\leq_2 \mathbf{T}^\#$. $\mathbf{T}^{lfp,\sigma 2} \not\leq_2 \mathbf{T}^*$. $\mathbf{T}^{lfp,\sigma 2} \not\leq_2 \mathbf{T}^\#$.

- $\mathbf{T}^c \not\leq_3 \mathbf{T}^*$. $\mathbf{T}^c \not\leq_3 \mathbf{T}^\#$. $\mathbf{T}^c \not\leq_2 \mathbf{T}^*$. $\mathbf{T}^c \not\leq_2 \mathbf{T}^\#$.
- $\mathbf{T}^{lfp,\sigma} \not\leq_3 \mathbf{T}^{lfp,\sigma}$. $\mathbf{T}^{lfp,\sigma^1} \not\leq_3 \mathbf{T}^{lfp,\sigma^1}$. $\mathbf{T}^{lfp,\sigma^2} \not\leq_2 \mathbf{T}^{lfp,\sigma^1}$.
- $\mathbf{T}^{gifp,\mu} \not\leq_2 \mathbf{T}^{gifp,\kappa}$. $\mathbf{T}^{gifp,\sigma} \not\leq_2 \mathbf{T}^{gifp,\sigma^1}$. $\mathbf{T}^{gifp,\sigma^2} \not\leq_2 \mathbf{T}^c$.
- $\mathbf{T}^{lfp,\kappa} \not\leq_1 \mathbf{T}^{lfp,\mu}$. $\mathbf{T}^{lfp,\rho'} \not\leq_1 \mathbf{T}^{gifp,\rho}$, where ρ and ρ' are chosen with ρ strictly to the left of ρ' from the list $\mu, \kappa, \sigma, \sigma^1, \sigma^2$.
- $\mathbf{T}^* \not\leq_1 \mathbf{T}^{gifp,\sigma}$. $\mathbf{T}^* \not\leq_1 \mathbf{T}^{gifp,\sigma^1}$. $\mathbf{T}^* \not\leq_1 \mathbf{T}^{gifp,\sigma^2}$.

We add the following three negative results:

- $\mathbf{T}^* \not\leq_2 \mathbf{T}^{lfp,\sigma}$. See [3], Example 6B.7.
- $\mathbf{T}^* \not\leq_2 \mathbf{T}^{lfp,\sigma^2}$. See [3], Example 6B.13.
- $\mathbf{T}^{lfp,\sigma} \not\leq_2 \mathbf{T}^{lfp,\kappa}$. Choose any S -neutral ground model. By Corollary 4.26, $lfp\sigma_M$ is classical. But $lfp\kappa_M$ is not classical, by the proof of Theorem 4.5.

The negative parts of Theorems 4.2, 4.5 and 4.21 follow from these results, together with (1) Lemma 4.20, (2) the fact that if $\mathbf{T} \leq_1 \mathbf{T}'$ then $\mathbf{T} \leq_2 \mathbf{T}'$, (3) the fact that \leq_1 and \leq_2 are reflexive and transitive, (4) the fact that \leq_3 is transitive, and (5) the positive parts of Theorems 4.2, 4.5 and 4.21.

Acknowledgements Many thanks to Anil Gupta. He answered some difficult questions with his Example 5.8 and his subtle Examples 5.10 and 5.11. Thanks to Michael Kremer for helpful conversations concerning both formal and methodological issues. Thanks also to an anonymous referee for helpful comments on an earlier draft.

References

1. Davis, L. (1979). An alternate formulation of Kripke's theory of truth. *Journal of Philosophical Logic*, 8, 289–296.
2. Grover, D. (1977). Inheritors and paradox. *Journal of Philosophy*, 84, 590–604.
3. Gupta, A., & Belnap, N. (1993). *The revision theory of truth*. Cambridge: MIT.
4. Haack, S. (1978). *Philosophy of logics*. Cambridge: Cambridge University Press.
5. Kirkham, R. (1992). *Theories of truth: a critical introduction*. Cambridge: MIT.
6. Kremer, M. (1988). Kripke and the logic of truth. *Journal of Philosophical Logic*, 17, 225–278.
7. Kremer, P. (2009). How truth behaves when there's no vicious reference, manuscript. <http://individual.utoronto.ca/philipkremer/onlinepapers/truthnvr.pdf>.
8. Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, 72, 690–716.
9. Kroon, F. (1975). Steinus on the paradoxes. *Theoria*, 50, 178–211.
10. Martin, R. L., & Woodruff, P. W. (1975). On representing 'True-in-L' in L. *Philosophia*, 5, 217–221.
11. Parsons, T. (1984). Assertion, denial and the liar paradox. *Journal of Philosophical Logic*, 13, 137–152.
12. Read, S. (1994). *Thinking about logic: An introduction to the philosophy of logic*. Oxford: Oxford University Press.
13. Visser, A. (1984). Four valued semantics and the liar. *Journal of Philosophical Logic*, 13, 181–212.
14. Visser, A. (1989). Semantics and the liar paradox. In D. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic* (Vol. 4). Dordrecht: Kluwer.
15. Woodruff, P. W. (1984). Paradox, truth and logic I. *Journal of Philosophical Logic*, 13, 213–232.