Minimalism's Constructions

Abstract

In this paper I introduce Horwich's deflationary theory of truth, called 'Minimalism', and I present his proposal to cope with the Liar paradox. This proposal proceeds by restricting the T-schema and, as a consequence of that, it needs a constructive specification of which instances of the T-schema are to be excluded from the minimalist theory of truth. Horwich has presented, in an informal way, one construction that specifies the minimalist theory. The main aim of the paper is to present and scrutinize some formal versions of Horwich's construction. I will also present a way of understanding these constructions as epistemic models of the relation of explanatory dependence between truth ascriptions and the extra-semantic facts.

Horwich has presented and defended his theory of truth in a number of places (see, specially, Horwich 1998, 2001, 2010b). Such theory, which is called 'the minimalist theory of truth', contains as axioms all instances of the T-schema applied to propositions. One of the main theses of Minimalism is that the instances of the T-schema are conceptually, explanatory and epistemologically fundamental.

As it is well known, though, the proposition that asserts its own untruth (let us call it 'the Liar') makes the theory consisting of just all instances of the T-schema inconsistent in classical logic (the argument to prove such inconsistency is the Liar Paradox). Until recently, Horwich's response to this problem was very succinct. In his (1998) he claims that the lesson the Liar tells us is that not all the instances of the T-schema are to be included as axioms in the theory (Horwich 1998, page 42). Consequently, the minimalist theory of truth must consist of a restricted collection of instances of the T-schema; only those that do not engender Liar-like paradoxes. Which of the instances of the T-schema should be removed, though, was left undetermined.

INorder to cope with the LIar, Horwich has presented, in his (2010a), a construction which offers a constructive specification of the instances of the T-schema we must exclude from the minimalist theory of truth. Horwich seems to defend a construction similar to the one proposed in Kripke (1975) and take the grounded propositions as the ones whose instances of the T-schema eventually constitute the minimalist theory of truth.

I will follow the formalisation of Horwich's construction as presented in Schindler (2015). For perspicuity, let us suppose we have a classical first-order language $\mathcal{L}$, the base language, and an expanded language $\mathcal{L}^+ = \mathcal{L} \cup \{Tr\}$ with a monadic predicate Tr intended to represent truth and suppose, furthermore, that for every formula $\phi \in \mathcal{L}^+$ we can express its canonical name $\ulcorner\phi\urcorner$ in $\mathcal{L}$ via some codification. I will suppose that $\mathcal{L}$ is strong enough to prove the Diagonal Lemma, so that the Liar paradox (and other paradoxes in the vicinity) can be formulated.

Given a model for the base language, $\mathcal{N}$, with domain D, I will use $\langle\mathcal{N}, A\rangle$ to refer to the model of the expanded language $\mathcal{L}^+$ whose interpretation of Tr is A, which will be a set of (codes of) sentences of $\mathcal{L}^+$. I will use $|\alpha|_{\mathcal{M}} = 1$ to mean that the formula α has semantic value 1 in the model $\mathcal{M}$ (and the same for having semantic value 0). Given a set of formulas Γ, I will use $|\Gamma|_{\mathcal{M}} = 1$ to mean that, for every $\gamma \in \Gamma$, $|\gamma|_{\mathcal{M}} = 1$. $\overline{D}$ is the set of (codes of) sentences of $\mathcal{L}^+$.

Let us begin with the construction. It will consist of a series H_σ of sets of sentences of $\mathcal{L}^+$ defined for every ordinal σ and relative to a model $\mathcal{N}$ for the base language. We need, first, the following definitions.

Definition

For any set A of formulas of $\mathcal{L}^+$, $A^- = \{\phi : \neg\phi \in A\}$.

For any $\phi \in \mathcal{L}^+$, T_ϕ is the ϕ-instance of the T-schema, i.e. $Tr\ulcorner\phi\urcorner \leftrightarrow \phi$.

For any set A of sentences of $\mathcal{L}^+$, $T_A = \{T_\phi : \phi \in A \text{ or } \phi \in A^-\}$.

Horwich's construction is formalised by the following series of sets of sentences of $\mathcal{L}^+$, given a model $\mathcal{N}$ for the base language and for any ordinal σ:

$$H_0 = \{\phi \in \mathcal{L} : |\phi|_{\mathcal{N}} = 1\}$$
$$H_{\sigma+1} = \{\phi \in \mathcal{L}^+ : H_\sigma \cup T_{H_\sigma} \models \phi\}$$
$$H_\lambda = \bigcup_{\alpha<\lambda} H_\alpha$$

where λ is a limit ordinal and $\models$ is implication in classical first-order logic.

This construction has a fixed point, that is an ordinal τ such that $H_\tau = H_{\tau+1}$. I will call the fixed point of the construction $\boldsymbol{H}$. Hence, at this point, Horwich's theory of truth, the minimalist theory of truth, is $T_{\boldsymbol{H}}$.

I will motivate two more constructions analogous to this one, the first is due to Schindler (2017) and uses the ω-rule to create the fixed point $\boldsymbol{H}^{\omega}$, which can be shown to be a subset of the fixed point we obtain with Kripke (1975)'s construction using the supervaluationsit scheme. I will cal this fixed point $\boldsymbol{VF}$. The second fixed point is constructed by treating the expressions in the base language $\mathcal{L}$ (the language without the truth predicate). I will cal this fixed point $\boldsymbol{H}^{*}$. It can be shown that $\boldsymbol{H}^{*}$ is just $\boldsymbol{VF}$.

It is well known that $\boldsymbol{VF}$ has many unsatisfactory properties. Specifically, the following statements are not in it:

1. For any sentence x; $Tr\ulcorner\neg\neg x\urcorner$ if, and only if, $Tr\ulcorner x\urcorner$.

2. For any sentences x, y; $Tr\ulcorner x \vee y\urcorner$ if, and only if, $Tr\ulcorner x\urcorner$ or $Tr\ulcorner y\urcorner$.

3. For any sentences x, y; $Tr\ulcorner x \wedge y\urcorner$ if, and only if, $Tr\ulcorner x\urcorner$ and $Tr\ulcorner y\urcorner$.

4. For any sentences x, y; if $Tr\ulcorner x \rightarrow y\urcorner$ and $Tr\ulcorner x\urcorner$, then $Tr\ulcorner y\urcorner$.

The problem is that we can find some instances that are not in it. Hence, they will neither be in $\boldsymbol{H}^{*}$, which makes it unsuitable for an appropriate theory of truth. This situation can be ameliorated if we restrict our already restricted logical consequence relation $\models_{\mathcal{N}}$ to models that have a maximally consistent set as extension of the truth predicate. The foxed point we obtain thus is $\boldsymbol{H}^{mc}$ which much stronger and it can be shown to contain principles 1-4 above.

Horwich (2010a) presents his construction as one that "squares with minimalism" (Horwich 2010a, p. 92, fn. 12) in the sense that it does not use compositional principles for truth, which are seen as incompatible with minimalism. This is so because Minimalism understands truth via the T-schema and not via compositional principles *à la* Tarski. Indeed, Horwich rejects Kripke's construction based on the strong Kleene scheme because it "invokes Tarski-style compositional principles" (Horwich 2010a, p. 92, ft. 12). At this point, then, it is natural to expect that Horwich would also reject the supervaluationist version of the Kripke's construction for, although he supervaluational scheme assigns the semantic value on the grounds of ways of making the truth predicate precise, in each of this ways, the semantic value of the sentences is achieved by compositional rules.[1] Importantly, this

[1] It could be said that the construction for $\boldsymbol{H}^{*}$ implicitly uses compositional principles for truth in the relation of logical consequence and that, in consequence, Horwich should reject it too. This is not the case, though, for Horwich thinks that truth does not play any role in the foundations of logic (see Horwich 1998, pages 74-6).

discussion clearly suggests that Horwich deos not take the construction to be a mere technical device to determine which instances of the T-schema are in the minimalist truth theory and that, hence, it needs to be under strict deflationist constraints—after all, if it were a mere technical device it would not matter whether it used compositional principles for truth.

Horwich's construction, though, heavily relies on the notion of groundedness; he himself claims that "a good solution to the liar paradox should articulate 'grounding' constraints [...] on which particular instances of [the T-schema] are axioms" (Horwich 2010a, page 91). But then, given that, as we concluded, the construction is not a mere technical device and has to satisfy strict deflationist constraints, using the notion of groundedness might seem to be at odds with Horwich's deflationist view about truth. After all, depending on how we understand the notion of groundedness, if it is constitutive of the notion of truth, it is no longer the case that "our commitment to [the T-schema] accounts for everything we do with the truth predicate" (Horwich 1998, page 121) and, hence, it is no longer the case that the T-schema implicitly defines it.

Horwich may have a way out of this situation; he might understand the construction of $\boldsymbol{H}^{mc}$ ($\boldsymbol{H}$, $\boldsymbol{H}^{\omega}$,$\boldsymbol{H}^{*}$) as a model of how truth claims are explained by some things in the world being in a certain way.[2] Let me elaborate on that. Horwich admits that the correspondence intuition can be accommodated into Minimalism. We can loosely characterize correspondence theories of truth as defending that being true consists in corresponding to facts. Although Horwich does not endorse, obviously, this characterization, he claims:

> [W]e might hope to accommodate much of what the correspondence theorist wishes to say without retreating an inch from our deflationary position. [...]
>
> It is indeed undeniable that whenever a proposition or an utterance is true, it is true *because* something in the world is a certain way—something typically external to the proposition or the utterance. For example,[...]
>
> $<$Snow is white$>$ is true *because* snow is white.
>
> (Horwich 1998, page 104)

[2] Thanks to ****** and ******* for their helpful insights on this issue.

Thus, claims Horwich, the fact that snow is white is explanatorily prior to the fact that <Snow is white> is true, in the same way that the basic laws of nature and the initial state of the universe are explanatorily prior to the fact that snow is white.

I think that it is reasonable to understand the construction of the fixed points we have discussed in this paper as an epistemic model of the relation of explanatory dependence between truth ascriptions and the extra-semantic facts. In this epistemic reading of groundedness, the grounded sentences, those in the fixed point, are the ones that can be explained—and, hence, the ones that can be known—given how the world is and given the appropriate instances of the T-schema.

References

Horwich, Paul (1998). *Truth*. 2nd. Oxford: Oxford University Press.

— (2001). "A Defense of Minimalism". In: *Synthese* 126.1-2, pages 149–65. Reprinted in Paul Horwich (2010b). *Truth-Meaning-Reality*. Oxford: Oxford University Press, pages 35–56.

— (2010a). "A Minimalist Critique of Tarski". In: *Truth-Meaning-Reality*. Oxford: Oxford University Press, pages 79–97.

— (2010b). *Truth-Meaning-Reality*. Oxford: Oxford University Press.

Kripke, Saul A. (1975). "Outline of a Theory of Truth". In: *Journal of Philosophy* 72.19, pages 690–716.

Schindler, Thomas (2015). "A Disquotational Theory of Truth as Strong as Z_2^-". In: *Journal of Philosophical Logic* 44, pages 395–410.

— (2017). "A Remark on Horwich's Notion of Grounding". Manuscript.