On Kripke's theory of truth — I

Stanislav Speranski

Steklov Mathematical Institute, Moscow

(SIMC)

20 May 2022

Kripke's Theory of Truth

Consider the signature of Peano arithmetic and its expansion obtained by adding an extra unary predicate symbol T, viz.

$$\sigma \; := \; \{0, \mathrm{s}, +, \times, =\} \quad \text{and} \quad \sigma_T \; := \; \sigma \cup \{T\}.$$

Throughout this presentation the following assumptions are in force:

- the connective symbols are $\neg$, $\wedge$ and $\vee$;
- the quantifier symbols are $\forall$ and $\exists$.

We abbreviate $\neg\varphi \vee \psi$ to $\varphi \to \psi$, $(\varphi \to \psi) \wedge (\psi \to \varphi)$ to $\varphi \leftrightarrow \psi$, etc. Let $\mathcal{L}$ and $\mathcal{L}_T$ be the first-order languages of σ and σ_T respectively.

Here is some related notation:

$$
\begin{aligned}
\mathit{For} &:= \text{the collection of all } \mathcal{L}\text{-formulas;}\\
\mathit{Sen} &:= \text{the collection of all } \mathcal{L}\text{-sentences;}\\
\mathit{For}_T &:= \text{the collection of all } \mathcal{L}_T\text{-formulas;}\\
\mathit{Sen}_T &:= \text{the collection of all } \mathcal{L}_T\text{-sentences.}
\end{aligned}
$$

Assume some Gödel numbering $\#$ of $\mathcal{L}_T$ has been chosen. Then we call $A \subseteq \mathbb{N}$ consistent iff there is no $\phi \in \mathit{Sen}_T$ s.t. both $\#\phi$ and $\#\neg\phi$ are in A. If $A \subseteq \mathbb{N}$, we write $\langle \mathfrak{N}, A \rangle$ for the expansion of the standard model $\mathfrak{N}$ of Peano arithmetic to σ_T in which T is interpreted as A.

In his 'Outline of a theory of truth', Kripke used partial interpretations of T, i.e. pairs of the form $S = \langle S^+, S^- \rangle$ where S^+ and S^- are disjoint subsets of $\mathbb{N}$, resp. called the extension of S and the anti-extension of S. Henceforth we limit ourselves to partial interpretations of T with consistent extensions. A partial valuation for σ_T is a mapping from Sen_T to a superset of $\{0, \frac{1}{2}, 1\}$.

By a valuation scheme we mean a function from partial interpretations to partial valuations. To begin with, let $\leqslant_{\mathrm{sK}}$ and $\leqslant_{\mathrm{wK}}$ be the orderings given by

$$0 \leqslant_{\mathrm{sK}} \tfrac{1}{2} \leqslant_{\mathrm{sK}} 1 \quad \text{and} \quad \tfrac{1}{2} \leqslant_{\mathrm{wK}} 0 \leqslant_{\mathrm{wK}} 1.$$

Define the strong Kleene valuation scheme V_{sK} inductively as follows:

- for any closed $\mathcal{L}$-terms t_1 and t_2,

$$V_{\mathrm{sK}}(S)(t_1 = t_2) := \begin{cases} 1 & \text{if } \mathfrak{N} \models t_1 = t_2, \\ 0 & \text{if } \mathfrak{N} \models t_1 \neq t_2; \end{cases}$$

- for every closed $\mathcal{L}$-term t,

$$V_{\mathrm{sK}}(S)(T(t)) := \begin{cases} 1 & \text{if } \langle \mathfrak{N}, S^+ \rangle \models T(t), \\ 0 & \text{if } \langle \mathfrak{N}, S^- \cup (\mathbb{N} \setminus \# Sen_T) \rangle \models T(t), \\ \frac{1}{2} & \text{otherwise}; \end{cases}$$

- $V_{\mathrm{sK}}(S)(\varphi \wedge \phi) := \min_{\leqslant_{\mathrm{sK}}} \{ V_{\mathrm{sK}}(S)(\varphi), V_{\mathrm{sK}}(S)(\phi) \}$;
- $V_{\mathrm{sK}}(S)(\forall x\, \varphi(x)) := \min_{\leqslant_{\mathrm{sK}}} \{ V_{\mathrm{sK}}(S)(\varphi(t)) \mid t \text{ is a closed } \mathcal{L}\text{-term} \}$;

- $V_{\mathrm{sK}}(S)(\varphi \vee \phi) := V_{\mathrm{sK}}(S)(\neg(\neg\varphi \wedge \neg\phi))$;
- $V_{\mathrm{sK}}(S)(\exists x\, \varphi(x)) := V_{\mathrm{sK}}(S)(\neg\forall x\, \neg\varphi(x))$;
- $V_{\mathrm{sK}}(S)(\neg\varphi) := 1 - V_{\mathrm{sK}}(S)(\varphi)$.

To get the weak Kleene valuation scheme V_{wK}, simply replace $\leqslant_{\mathrm{sK}}$ by $\leqslant_{\mathrm{wK}}$. Next we turn to so-called supervaluation schemes, each of which has the form

$$V(S)(\varphi) := \begin{cases} 1 & \text{if for all } A \subseteq \mathbb{N} \text{ satisfying } [*],\ \langle \mathfrak{N}, A\rangle \models \varphi, \\ 0 & \text{if for all } A \subseteq \mathbb{N} \text{ satisfying } [*],\ \langle \mathfrak{N}, A\rangle \models \neg\varphi, \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

The best known such schemes are V_{SV}, V_{VB}, V_{FV} and V_{MC}, given by:

$$\begin{aligned}
V &= V_{\mathrm{SV}} &\iff& [*] = \text{‘}S^+ \subseteq A\text{’};\\
V &= V_{\mathrm{VB}} &\iff& [*] = \text{‘}S^+ \subseteq A \text{ and } A \cap S^- = \varnothing\text{’};\\
V &= V_{\mathrm{FV}} &\iff& [*] = \text{‘}S^+ \subseteq A \text{ and } A \text{ is consistent’};\\
V &= V_{\mathrm{MC}} &\iff& [*] = \text{‘}S^+ \subseteq A \text{ and } A \text{ is cons. and complete’}.
\end{aligned}$$

Here the ‘completeness’ of A means that for each $\phi \in Sen_T$ we have $\#\phi \in A$ or $\#\neg\phi \in A$. The last scheme emerges from Leitgeb's ‘What truth depends on’ (although the definition presented below was stated explicitly by his PhD student Thomas Schindler). Say that $\varphi \in Sen_T$ depends on $A \subseteq \mathbb{N}$ iff for every $B \subseteq \mathbb{N}$,

$$\langle \mathfrak{N}, B \rangle \models \varphi \iff \langle \mathfrak{N}, B \cap A \rangle \models \varphi.$$

Now define Leitgeb's valuation scheme V_{L} by

$$V_{\mathrm{L}}(S)(\varphi) := \begin{cases} 1 & \text{if } \varphi \text{ depends on } S^+ \cup S^- \text{ and } \langle \mathfrak{N}, S^+ \rangle \models \varphi, \\ 0 & \text{if } \varphi \text{ depends on } S^+ \cup S^- \text{ and } \langle \mathfrak{N}, S^+ \rangle \models \neg\varphi, \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

It should be noted that each valuation scheme V induces a function $\mathfrak{J}_V$ from partial interpretations to partial interpretations, called the Kripke-jump operator for V, as follows:

$$\begin{aligned} \mathfrak{J}_V(S)^+ &:= \{\#\varphi \mid \varphi \in Sen_T \text{ and } V(S)(\varphi) = 1\}, \\ \mathfrak{J}_V(S)^- &:= \{\#\varphi \mid \varphi \in Sen_T \text{ and } V(S)(\varphi) = 0\} \cup \\ &\qquad \cup \{n \in \mathbb{N} \mid n \notin \#Sen_T\}. \end{aligned}$$

In turn $\mathcal{J}_V$ generates a transfinite sequence indexed by ordinals:

$$\mathcal{J}_V^\alpha(S) := \begin{cases} S & \text{if } \alpha = 0, \\ \mathcal{J}_V\left(\mathcal{J}_V^\beta(S)\right) & \text{if } \alpha = \beta + 1, \\ \langle \bigcup_{\beta<\alpha} \mathcal{J}_V^\beta(S)^+, \bigcup_{\beta<\alpha} \mathcal{J}_V^\beta(S)^- \rangle & \text{if } \alpha \in \text{L-Ord}. \end{cases}$$

We shall often write T_V^α instead of $\mathcal{J}_V^\alpha(\varnothing, \varnothing)^+$ — these sets constitute the truth hierarchy for V.

Moreover Kripke dealt with monotone schemes, i.e. those which satisfy the condition that for any partial interpretations S_1 and S_2,

$$S_1^+ \subseteq S_2^+ \;\&\; S_1^- \subseteq S_2^- \implies$$
$$\implies \mathcal{J}_V(S_1)^+ \subseteq \mathcal{J}_V(S_2)^+ \;\&\; \mathcal{J}_V(S)^- \subseteq \mathcal{J}_V(S)^-.$$

Observation (Kripke)

For every monotone valuation scheme V there exists an ordinal α s.t. $\mathcal{J}_V^{\alpha+1}(\varnothing, \varnothing) = \mathcal{J}_V^{\alpha}(\varnothing, \varnothing)$ — yielding the least fixed point of $\mathcal{J}_V$.

It is easy to verify that each $V \in \{V_{\mathrm{sK}}, V_{\mathrm{wK}}, V_{\mathrm{SV}}, V_{\mathrm{VB}}, V_{\mathrm{FV}}, V_{\mathrm{MC}}, V_{\mathrm{L}}\}$ is monotone and furthermore has the following properties:

- if $\mathcal{J}_V(S) = S$, then $V(S)(T(\ulcorner\varphi\urcorner)) = V(S)(\varphi)$;
- $\#\varphi \in \mathcal{J}_V^{\alpha}(S)^-$ iff $\#\neg\varphi \in \mathcal{J}_V^{\alpha}(S)^+$;
- $\#\varphi \in \mathcal{J}_V^{\alpha}(S)^+$ iff $\#\neg\varphi \in \mathcal{J}_V^{\alpha}(S)^-$;
- $\mathcal{J}_V$ turns out to be a 'Π_1^1-operator' — so by a well-known theorem of Spector, $\mathrm{T}_V^{\alpha} = \mathrm{T}_V^{\alpha+1}$ already for some $\alpha \in$ C-Ord $\cup \{\omega_1^{\mathrm{CK}}\}$.

- J. P. Burgess (1986). The truth is never simple. *Journal of Symbolic Logic* 51:4, 663–681.
- J. Cain and Z. Damnjanovic (1991). On the weak Kleene scheme in Kripke's theory of truth. *Journal of Symbolic Logic* 56:4, 1452–1468.
- S. Feferman (2008). Axioms for determinateness and truth. *Review of Symbolic Logic* 1:2, 204–217.
- M. Fischer, V. Halbach, J. Kriener and J. Stern (2015). Axiomatizing semantic theories of truth? *Review of Symbolic Logic* 8:2, 257–278.
- S. Kripke (1975). Outline of a theory of truth. *The Journal of Philosophy* 72:19, 690–716.

- H. Leitgeb (2005). What truth depends on. *Journal of Philosophical Logic* 34:2, 155–192.
- T. Schindler (2015). *Type-Free Truth* (Ph.D. Thesis). Ludwig-Maximilians-Universität München.
- S. O. Speranski (2017). Notes on the computational aspects of Kripke's theory of truth. *Studia Logica* 105:2, 407–429. ✓
- P. D. Welch (2014). The complexity of the dependence operator. *Journal of Philosophical Logic* 44:3, 337–340.