Axiomatic theories of truth based on Weak and Strong Kleene logic

Cezary Cieśliński

Institute of Philosophy
University of Warsaw
Poland

Udine 2018

Outline

- The objective is to compare the conceptual strength of two axiomatic theories of truth: *KF* and *WKF*.
- *KF* is an axiomatic theory of truth designed to capture Kripke's fixed-point model construction based on the Strong Kleene evaluation schema.
- *WKF* is like *KF*, except that it is based on the Weak Kleene evaluation schema.
- The main result is that *WKF* is conceptually weaker than *KF*.

Comparing axiomatic theories of truth

Various axiomatic theories of truth have been proposed in the literature. One of the main research topics has been that of assessing their strength.

How to measure the strength of theories? Here are some basic options:

- inclusion (the simplest but also a very crude method),
- comparisons of arithmetical strength.

Recently Kentaro Fujimoto proposed a more subtle measure, which permits to compare *conceptual* strength of theories.

Truth-definability

Denoting by L_T the result of adding a new unary predicate '$T(x)$' to the language of first order arithmetic, we define:

Definition

Let Th_1 and Th_2 be theories in L_T. Th_1 is relatively truth-definable in Th_2 (or Th_2 defines the truth predicate of Th_1) iff there is a formula $\theta(x) \in L_T$ such that for every $\psi \in L_T$, if $Th_1 \vdash \psi$, then $Th_2 \vdash \psi(\theta(x)/T(x))$.

General idea:

If Th_2 defines the truth predicate of Th_1, then Th_2 is not *conceptually* weaker than Th_1. In such a situation Th_2 contains the resources permitting us to reproduce the very concept of truth characterised by the axioms of Th_1.

Kripke-Feferman system

Definition (*KF*)

1 $\forall s, t \in Tm^c\big(T(s = t) \equiv val(s) = val(t)\big)$

2 $\forall s, t \in Tm^c\big(T(\neg s = t) \equiv val(s) \neq val(t)\big)$

3 $\forall \varphi\big(Sent_{L_T}(\varphi) \rightarrow (T(\neg\neg\varphi) \equiv T\varphi)\big)$

4 $\forall \varphi \forall \psi\big(Sent_{L_T}(\varphi \wedge \psi) \rightarrow (T(\varphi \wedge \psi) \equiv T\varphi \wedge T\psi)\big)$

5 $\forall \varphi \forall \psi\big(Sent_{L_T}(\varphi \wedge \psi) \rightarrow (T\neg(\varphi \wedge \psi) \equiv T\neg\varphi \vee T\neg\psi)\big)$

6-7 Similarly for disjunction

8 $\forall v \forall \varphi(x)\big(Sent_{L_T}(\forall v \varphi(v)) \rightarrow (T(\forall v \varphi(v)) \equiv \forall t \in Tm^c T(\varphi(t)))\big)$

9 $\forall v \forall \varphi(x)\big(Sent_{L_T}(\forall v \varphi(v)) \rightarrow (T(\neg\forall v \varphi(v)) \equiv \exists t \in Tm^c T(\neg\varphi(t)))\big)$

10-11 Similarly for the existential quantifier

12 $\forall t \in Tm^c\big(T(Tt) \equiv T(val(t))\big)$

13 $\forall t \in Tm^c\big(T\neg Tt \equiv (T(\neg val(t)) \vee \neg Sent_{L_T}(val(t)))\big)$

Weak Kripke-Feferman

We adopt the following abbreviations:

- Let $\varphi \in Sent_{L_T}$. Then '$D(\varphi)$' ('x is determined') is the formula: '$T(\varphi) \vee T(\neg\varphi)$'.
- Let $\varphi(x)$ be a formula of L_T with one variable free. Then '$D(\varphi(x))$' is the formula: '$\forall x\, \big[T(\varphi(x)) \vee T(\neg\varphi(x))\big]$'.
- The expression '$D(\varphi, \psi)$' abbreviates '$D(\varphi) \wedge D(\psi)$'.

Below only those axioms of *WKF* are listed which differ from the corresponding axioms of *KF*.

5 $\forall \varphi, \psi \in Sent_{L_T}\big[T\big(\neg(\varphi \wedge \psi)\big) \equiv \big(D(\varphi, \psi) \wedge (T(\neg\varphi) \vee T(\neg\psi))\big)\big]$

6 $\forall \varphi, \psi \in Sent_{L_T}\big[T(\varphi \vee \psi) \equiv \big(D(\varphi, \psi) \wedge (T(\varphi) \vee T(\psi))\big)\big]$

9 $\forall \varphi(x) \in L_T \forall v \in Var\big[T(\neg\forall v \varphi(v)) \equiv \big(D(\varphi(x)) \wedge \exists t T(\neg\varphi(t))\big)\big]$

10 $\forall \varphi(x) \in L_T \forall v \in Var\big[T(\exists v \varphi(v)) \equiv \big(D(\varphi(x)) \wedge \exists t T(\varphi(t))\big)\big]$

Weak and strong Kleene Jump

Definition (Strong Kleene Jump)

Let $S \subseteq \omega$. We define:

$$
\begin{aligned}
J^{SK}(S) = \ & \{\ulcorner t = s \urcorner : val(t) = val(s)\} \\
\cup \ & \{\ulcorner t \neq s \urcorner : val(t) \neq val(s)\} \\
\cup \ & \{\ulcorner T(t) \urcorner : val(t) \in S\} \\
\cup \ & \{\ulcorner \neg T(t) \urcorner : \neg val(t) \in S \vee \neg Sent_{L_T}(val(t))\} \\
\cup \ & \{\ulcorner \varphi \circ \psi \urcorner : \varphi \in S \circ \psi \in S\} \\
\cup \ & \{\ulcorner \neg(\varphi \circ \psi) \urcorner : \neg\varphi \in S \circ_d \neg\psi \in S\} \\
\cup \ & \{\ulcorner Qv\varphi \urcorner : Qt \in Tm^c(\varphi(t) \in S)\} \\
\cup \ & \{\ulcorner \neg Qv\varphi \urcorner : Q_d t \in Tm^c(\neg\varphi(t) \in S)\}.
\end{aligned}
$$

Weak and strong Kleene Jump

Let $S \subseteq \omega$. Let '$D(S, \varphi)$' be a shorthand for: '$\varphi \in S \vee \neg\varphi \in S$'. Let '$D(S, \varphi(x))$' abbreviate '$\forall t \in Tm^c(\varphi(t) \in S \vee \neg\varphi(t) \in S)$'. Let '$D(S, \varphi, \psi)$' be a shorthand for '$D(S, \varphi) \wedge D(S, \psi)$'.

Definition (Weak Kleene Jump)

$$
\begin{aligned}
J^{WK}(S) = \ & \{\ulcorner t = s \urcorner : val(t) = val(s)\} \\
\cup \ & \{\ulcorner t \neq s \urcorner : val(t) \neq val(s)\} \\
\cup \ & \{\ulcorner T(t) \urcorner : val(t) \in S\} \\
\cup \ & \{\ulcorner \neg T(t) \urcorner : \neg val(t) \in S \vee \neg Sent_{L_T}(val(t))\} \\
\cup \ & \{\ulcorner \varphi \circ \psi \urcorner : D(S, \varphi, \psi) \wedge (\varphi \in S \circ \psi \in S)\} \\
\cup \ & \{\ulcorner \neg(\varphi \circ \psi) \urcorner : D(S, \varphi, \psi) \wedge (\neg\varphi \in S \circ_d \neg\psi \in S)\} \\
\cup \ & \{\ulcorner Qv\varphi \urcorner : D(S, \varphi(v)) \wedge Qt \in Tm^c(\varphi(t) \in S)\} \\
\cup \ & \{\ulcorner \neg Qv\varphi \urcorner : D(S, \varphi(v)) \wedge Q_d t \in Tm^c(\neg\varphi(t) \in S)\}.
\end{aligned}
$$

Models of *KF* and *WKF*

Definition

Let E be either Weak Kleene or Strong Kleene evaluation scheme.

- $T^E_0 = Th(N)$,
- $T^E_{\alpha+1} = J^E(T^E_\alpha)$,
- $T^E_\lambda = \bigcup_{\alpha<\lambda} T^E_\alpha$,
- T^E is T^E_κ for the least κ such that $T^E_\kappa = T^E_{\kappa+1}$.

(N, T^{SK}) and (N, T^{WK}) are the least fixed-point models based on Strong Kleene and Weak Kleene logic. They are also models for *KF* and *WKF*, respectively.

Some known results

Definition

$ClOrd_{(N,T^E)}$ is the least κ such that $T^E_\kappa = T^E_{\kappa+1}$.

Theorem

(a) $ClOrd_{(N,T^{SK})} = \omega_1^{CK}$ *(the Church-Kleene ordinal),*

(b) $ClOrd_{(N,T^{WK})}$ *can be* ω *or* ω_1^{CK} *or infinitely many ordinals in between.*

Source:

(a) Kripke, S. 'Outline of a theory of truth', *The Journal of Philosophy* 1975,

(b) Cain, J. and Damnjanovic, Z. 'On the Weak Kleene Scheme in Kripke's Theory of Truth', *The Journal of Symbolic Logic* 1991.

Some known results

Theorem

(a) KF and WKF have the same arithmetical strength, given that their base arithmetical theory is PA formulated in the language of PRA.

(b) KF defines the truth predicate of WKF.

Source: Fujimoto K. ‘Relative truth definability of axiomatic truth theories’, *Bulletin of Symbolic Logic* 2010.

Question:

Does *WKF* define the truth predicate of *KF*?

Main lemmas

Definition

Let E be either Weak Kleene or Strong Kleene evaluation scheme. A sentence $\psi \in L_T$ is E-grounded iff either ψ or $\neg\psi$ belongs to T^E.

Lemma 1

There are no formulas $\tau(x)$ and $G(x) \in L_T$ such that $\tau(x)$ is a *KF* truth predicate in (N, T^{WK}) and for every $\psi \in Sent_{L_T}$:

(a) ψ is WK-grounded iff $(N, T^{WK}) \models \tau(G(\psi))$,

(b) ψ is WK-ungrounded iff $(N, T^{WK}) \models \tau(\neg G(\psi))$.

Lemma 2

If (N, T^{WK}) has a truth predicate of *KF*, then there are formulas $\tau(x)$ and $G(x)$ satisfying the conditions (a) and (b) from Lemma 1.

Proof of Lemma 1 (idea)

Lemma 1 is proved in the following way. Assuming that the lemma is false, define:

$$Tr(x) := \tau(x[G(t) \wedge T(t)/T(t)]).$$

In other words, $Tr(x)$ is the formula stating that the result of substituting '$G(t) \wedge T(t)$' for every occurrence of '$T(t)$' in x satisfies τ.

The proof is completed by showing that then:

$$\text{For every } \psi \in Sent_{L_T}, (N, T^{WK}) \models Tr(\psi) \equiv \psi,$$

which contradicts Tarski's undefinability theorem. The argument proceeds by induction on the complexity of ψ. For ψ of the form $T(t)$ or $\neg T(t)$, the conditions (a) and (b) from Lemma 1 are crucially used.

Lemma 2 - preliminaries

Definition

- $x \lhd y$ is an abbreviation of the following arithmetical formula:
  $Sent_{L_T}(x) \wedge Sent_{L_T}(y) \wedge$
  $\Big(\exists t \in Tm^c(y = \ulcorner T(t)\urcorner \wedge x = val(t))$
  $\vee\ \exists \psi \in Sent_{L_T}(y = \ulcorner \neg\psi \urcorner \wedge x = \psi)$
  $\vee\ \exists \varphi, \psi \in Sent_{L_T}(y = \ulcorner \varphi \circ \psi \urcorner \wedge x = \varphi \vee x = \varphi)$
  $\vee\ \exists \theta(x) \in Fm_T^{\leq 1} \exists t \in Tm^c \exists v \in Var(y = \ulcorner Qv\theta(v)\urcorner \wedge x = \ulcorner \theta(t)\urcorner)\Big)$.
- $\lhd^*$ denotes the transitive closure of $\lhd$,
- a sentence $\psi \in L_T$ is well-founded iff $\lhd$ is well-founded on $\{x : x \unlhd^* \psi\}$.

Observation

For every sentence $\psi \in L_T$(ψ is WK-grounded iff ψ is well-founded).

Proof of Lemma 2: main cases

For the proof of Lemma 2, we assume that $\tau(x)$ defines a truth predicate of KF in (N, T^{WK}). Our task is to construct formulas $\tau'(x)$ and $G(x)$ such that $G(x)$ expresses (in the scope of $\tau'(x)$) the property of WK-groundedness. The proof proceeds by analysis cases.

Cases

Case 1: $ClOrd_{(N,T^{WK})} < \omega_1^{CK}$,

Case 2: $ClOrd_{(N,T^{WK})} = \omega_1^{CK}$.

Remark: the above two cases cover the choice of coding and the choice of base arithmetical language.

Case 1: $ClOrd_{(N,T^{WK})} < \omega_1^{CK}$

For an arbitrary well-founded ψ define:

$$Ord(\psi) = sup\{Ord(x) : x \lhd \psi\}.$$

Observation

$\exists\kappa < \omega_1^{CK}\ \forall\psi \in Sent_{L_T}\big(\psi \textit{ is well-founded} \to Ord(\psi) < \kappa\big).$

Theorem (Kripke 1975)

Let S be an arbitrary set of natural numbers. The following conditions are equivalent:

- *S is* Δ^1_1*,*
- *there is a formula* $\varphi(x) \in L_T$ *such that* $\varphi(x)$ *is total in* (N, T^{SK}) *and* $S = \{n : (N, T^{SK}) \models T(\varphi(n))\}$*.*

Case 1: $ClOrd_{(N,T^{WK})} < \omega_1^{CK}$

Lemma 2 follows from the following observation.

Observation

The set of sentences well-founded in (N, T^{WK}) *is* Δ^1_1.

Proof.

The Π^1_1 formulation is straightforward. For a Σ^1_1 formulation, let κ be an ordinal smaller that ω_1^{CK} such that $Ord(\psi) < \kappa$ for every well-founded ψ. Then the well-foundedness of an arbitrary ψ can be expressed by means of the following Σ^1_1 formula:

$$\exists f \exists \alpha < \kappa \big(f \text{ is a surjection mapping } \{x : x \trianglelefteq^* \psi\} \text{ onto } \alpha \wedge \forall x, y \trianglelefteq^* \psi (x \triangleleft y \rightarrow f(x) < f(y))\big).$$

Note that since $\kappa < \omega_1^{CK}$, we can treat the quantification over ordinals as quantification over ordinal notations. □

Case 2: $ClOrd_{(N,T^{WK})} = \omega_1^{CK}$

- In this case, the reasoning from Case 1 is inapplicable, since we cannot restrict the ordinals of well-founded formulas by any ordinal below ω_1^{CK}.
- The proof of Lemma 2 is now based on the insight that well-founded sentences form a structure which is complex enough to permit us (in the presence of the *KF* truth predicate) to reconstruct in (N, T^{WK}) various Strong Kleene model-theoretic constructions, including the Kripkean construction of the least fixed-point model.
- In short: we are going to treat well-founded sentences as ordinals.

Definition

Let ψ be the diagonal formula satisfying (provably in PAT) the condition:

$$
\begin{aligned}
\psi(s,x) \equiv\ & s \in Sent_{L_T} \wedge \\
& \Big(x \in Sent_{L_{PA}} \wedge T(x) \\
& \vee\ x = \ulcorner T(t) \urcorner \wedge \exists s' \lhd s T(\psi(s', val(t))) \\
& \vee\ x = \ulcorner \neg T(t) \urcorner \wedge \big(\exists s' \lhd s T(\psi(s', \neg val(t))) \vee \neg Sent_{L_T}(val(t))\big) \\
& \vee\ x = \ulcorner \neg\neg\varphi \urcorner \wedge \exists s' \lhd s T(\psi(s', \varphi)) \\
& \vee\ x = \ulcorner \varphi \circ \chi \urcorner \wedge \exists s' s'' \lhd s \big(T(\psi(s', \varphi)) \circ T(\psi(s'', \chi))\big) \\
& \vee\ x = \ulcorner \neg(\varphi \circ \chi) \urcorner \wedge \exists s' s'' \lhd s \big(T(\psi(s', \neg\varphi)) \circ_d T(\psi(s'', \neg\chi))\big) \\
& \vee\ x = \ulcorner Qv\varphi \urcorner \wedge \exists s' \lhd s QaT(\psi(s', \varphi(a))) \\
& \vee\ x = \ulcorner \neg Qv\varphi \urcorner \wedge \exists s' \lhd s Q_d a T(\psi(s', \neg\varphi(a)))\Big)
\end{aligned}
$$

Definability of Strong Kleene least fixed point construction

The next lemma establishes a connection between the behaviour of ψ under τ and Kripke's least fixed-point construction.

Lemma

$\forall s \forall \alpha [Ord(s) = \alpha \rightarrow \forall \varphi \in Sent_{L_T} (\varphi \in T^{SK}_{\alpha} \equiv (N, T^{WK}) \models \tau(\psi(s, \varphi)))]$.

Corollary

If $\tau(x)$ is an arbitrary KF truth predicate in (N, T^{WK}), then there is a formula $\tau'(x) \in L_T$ which defines in (N, T^{WK}) the set of sentences determined as true in the least fixed point model of KF.

Proof.

Abbreviating $T(s) \vee T(\neg s)$ as $s \in D$, define $\tau'(x)$ as '$\exists s \in D\ \tau(\psi(s, x))$'. The result follows immediately from the above Lemma together with the observation that the ordinals of well-founded sentences are arbitrarily large below ω_1^{CK}. □

Something is still missing

We need to show the existence of a *KF* truth predicate which can ‘recognize’ Weak Kleene groundedness. For this we need to define a formula $G(x)$ expressing groundedness.
By diagonal lemma, let $\theta(x)$ be a formula of L_T such that *PAT* proves that:

$$\theta(x) \equiv \exists y \lhd x \; T(\theta(y)).$$

Now, our $G(x)$ will be defined as $\neg\theta(x)$. Let us start with the following:

Observation

(a) $\forall \psi \in Sent_{L_T} \big(\psi$ *is WK-grounded iff* $(N, T^{SK}) \models T(\neg\theta(\psi))\big)$,
(b) $\neg\exists\psi \in Sent_{L_T}(N, T^{SK}) \models T(\theta(\psi))$.

Problem: the model ‘recognizes’ Weak Kleene groundedness only.

Adding $\theta(\psi)$ for all WK-ungrounded sentences ψ

Definition

- $T_0^\theta = Th(N)$,
- $T_{\alpha+1}^\theta = \{\theta(t) : t \in Tm^c \wedge val(t) \text{ is WK-ungrounded } \} \cup J^{SK}(T_\alpha^\theta)$,
- $T_\lambda^\theta = \bigcup_{\alpha<\lambda} T_\alpha^\theta$,
- T^θ is T_κ^θ for the least κ such that $T_\kappa^\theta = T_{\kappa+1}^\theta$.

Observation

$(N, T^\theta) \models KF$ *and for every* $\psi \in Sent_{L_T}$*:*

(a) ψ *is WK-grounded iff* $(N, T^\theta) \models T(\neg\theta(\psi))$*,*

(b) ψ *is WK-ungrounded iff* $(N, T^\theta) \models T(\theta(\psi))$*.*

In order to finish the proof of Lemma 2, it is enough to demonstrate that if (N, T^{WK}) defines a truth predicate of *KF*, then it defines also T^θ.

Assume that (N, T^{WK}) defines a truth predicate of KF. Let $\tau(x)$ be a formula defining in (N, T^{WK}) the set of sentences determined as true in the least fixed-point model of KF. Define:

$$\tau^{Compl}(x) := \neg\tau(\neg x).$$

Then $\tau^{Compl}(x)$ defines in (N, T^{WK}) the set of sentences determined as true in the largest fixed-point model of KF.

Definition

Let ψ_1 be such that:

$$
\begin{aligned}
\psi_1(s,x) \equiv\ & (x \in Sent_{L_{PA}} \wedge T(x)) \\
& \vee\ x = \ulcorner \theta(t) \urcorner \wedge T(\theta(t)) \\
& \vee\ x = \ulcorner T(t) \urcorner \wedge \exists s' \lhd s T(\psi_1(s', val(t))) \\
& \vee\ x = \ulcorner \neg T(t) \urcorner \wedge \big(\exists s' \lhd s T(\psi_1(s', \neg val(t))) \vee \neg Sent_{L_T}(val(t))\big) \\
& \vee\ x = \ulcorner \neg\neg\varphi \urcorner \wedge \exists s' \lhd s T(\psi_1(s', \varphi)) \\
& \vee\ x = \ulcorner \varphi \circ \chi \urcorner \wedge \exists s' s'' \lhd s \big(T(\psi_1(s', \varphi)) \circ T(\psi_1(s'', \chi))\big) \\
& \vee\ x = \ulcorner \neg(\varphi \circ \chi) \urcorner \wedge \exists s' s'' \lhd s \big(T(\psi_1(s', \neg\varphi)) \circ_d T(\psi_1(s'', \neg\chi))\big) \\
& \vee\ x = \ulcorner Qv\varphi \urcorner \wedge \exists s' \lhd s Qa T(\psi_1(s', \varphi(a))) \\
& \vee\ x = \ulcorner \neg Qv\varphi \urcorner \wedge \exists s' \lhd s Q_d a T(\psi_1(s', \neg\varphi(a)))
\end{aligned}
$$

Definability of T^θ in (N, T^{WK})

Finally, we define:

Definition

$\tau^\theta(x) := \exists s\big(D(s) \wedge \tau^{Compl}(\psi(s, x))\big)$.

Then $\tau^\theta(x)$ defines T^θ in (N, T^{WK}).

THE END

Thanks for your attention!!!