AXIOMATIC THEORIES OF TRUTH: A SURVEY

Volker Halbach

7th July 2022

13th Panhellenic Logic Symposium

Volos

Slides are here: `https://users.ox.ac.uk/~sfop0114/pdf/volos.pdf`

1. History
2. Axiomatic Theories of Truth in Philosophy
3. Classification of truth theories
4. Disquotational theories
5. Compositional theories
6. Conclusion

In an axiomatic theory of truth, the truth or satisfaction predicate is taken to be a primitive expression.

Tarski (1935) formulated and studied formal axiomatic truth theories. His adequacy condition can be understood as an axiomatization. However, in the end he rejected the axiomatic approach.

Davidson advocated axiomatic truth theories in the 1960s without formulating the theories.

In the late 1970s Feferman started to work on axiomatizations of Kripke's theory.

In an axiomatic theory of truth, the truth or satisfaction predicate is taken to be a primitive expression.

Tarski (1935) formulated and studied formal axiomatic truth theories. His adequacy condition can be understood as an axiomatization. However, in the end he rejected the axiomatic approach.

Davidson advocated axiomatic truth theories in the 1960s without formulating the theories.

In the late 1970s Feferman started to work on axiomatizations of Kripke's theory.

In an axiomatic theory of truth, the truth or satisfaction predicate is taken to be a primitive expression.

Tarski (1935) formulated and studied formal axiomatic truth theories. His adequacy condition can be understood as an axiomatization. However, in the end he rejected the axiomatic approach.

Davidson advocated axiomatic truth theories in the 1960s without formulating the theories.

In the late 1970s Feferman started to work on axiomatizations of Kripke's theory.

Truth and satisfaction predicates are frequently and essentially used in various areas of philosophy.

Examples

- S knows p iff S believes p, S is justified in his/her belief p and p is true (and some Gettier condition is satisfied).
- Whatever I clearly and distinctly perceive is true.
- There are unknowable truths. $\exists x\,(\neg \Diamond Kx \wedge Tx)$
- There are synthetic a priori truths.

Truth and satisfaction predicates are frequently and essentially used in various areas of philosophy.

Examples

- S knows p iff S believes p, S is justified in his/her belief p and p is true (and some Gettier condition is satisfied).
- Whatever I clearly and distinctly perceive is true.
- There are unknowable truths. $\exists x\, (\neg \Diamond Kx \wedge Tx)$
- There are synthetic a priori truths.

Truth and satisfaction predicates are frequently and essentially used in various areas of philosophy.

Examples

- *S* knows *p* iff *S* believes *p*, *S* is justified in his/her belief *p* and *p* is true (and some Gettier condition is satisfied).
- Whatever I clearly and distinctly perceive is true.
- There are unknowable truths. $\exists x\,(\neg \Diamond Kx \wedge Tx)$
- There are synthetic a priori truths.

Truth and satisfaction predicates are frequently and essentially used in various areas of philosophy.

Examples

- S knows p iff S believes p, S is justified in his/her belief p and p is true (and some Gettier condition is satisfied).
- Whatever I clearly and distinctly perceive is true.
- There are unknowable truths. $\exists x\,(\neg \Diamond Kx \wedge Tx)$
- There are synthetic a priori truths.

For philosophical applications we require a truth predicate that is not relativized to a model or structure.

The truth predicate is global in the sense that it is sensibly applicable to arbitrary sentences of one's language (or, perhaps, to all propositions), not just some sublanguage.

Philosophers usually see this, but are not bothered. They tell the logicians to go away and only come back once they have found a solution.

For philosophical applications we require a truth predicate that is not relativized to a model or structure.

The truth predicate is global in the sense that it is sensibly applicable to arbitrary sentences of one's language (or, perhaps, to all propositions), not just some sublanguage.

Philosophers usually see this, but are not bothered. They tell the logicians to go away and only come back once they have found a solution.

Nonclassical Logic

Don't trust the logicians!

How nonclassical logic spreads:

S knows p iff S believes p, S is justified in his/her belief p and p is true (and some Gettier condition is satisfied).

All sentences of the form $\varphi \to \varphi$ are analytic.

Nonclassical Logic

Don't trust the logicians!

How nonclassical logic spreads:

S knows p iff S believes p, S is justified in his/her belief p and p is true (and some Gettier condition is satisfied).

All sentences of the form $\varphi \rightarrow \varphi$ are analytic.

Nonclassical Logic

Don't trust the logicians!

How nonclassical logic spreads:

S knows p iff S believes p, S is justified in his/her belief p and p is true (and some Gettier condition is satisfied).

All sentences of the form $\varphi \rightarrow \varphi$ are analytic.

Reduction

A truth predicate can serve certain purposes in reductions. Adding a truth or satisfaction predicate to a language has effects that are similar to adding propositional quantifiers or second-order quantifiers.

Hence commitment to second-order objects can be eliminated by the use of a truth or satisfaction predicate.

1. non-classical theories

   There is much work on paraconsistent and other nonclassical 'logics of truth' but less on full theories of truth with a base theory Field (2008), Kremer (1988), Feferman (1984), Halbach and Horsten (2006), Leigh and Rathjen (2012)

2. classical theories

   They can be categorized along the following criteria:

   1. typed vs type-free
   2. disquotational vs compositional

Before we can formulate a theory of truth, we should have a theory of truth bearers, e.g., a theory of syntax (Halbach and Leigh 2022), propositions, or the like.

I'll use (first-order) Peano arithmetic with a suitable coding. Many of the results can be be applied to other base theories, e.g. a theory of concatenation, Zermelo-Fraenkel set theory, 'disentangled' theories.

So the truth theories are formulated in the language of first-order arithmetic plus a unary predicate T for truth.

All theories considered below are formulated in classical logic.

Plan

- Disquotational Theories
  - Typed Disquotational Theories
  - Untyped Disquotational Theories
- Compositional Theories
  - Typed Compositional Theories
  - Untyped Compositional Theories

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

1. Is truth definable? Is it reducible, conservative or eliminable?
2. Which axioms and rules about truth can be sensibly combined?
3. What's the expressive and deductive power of truth? What's the role of truth in reasoning?
4. To what extent can theories of truth replace second-order quantifiers and play a role in foundations?
5. How can truth be used to make explicit assumptions implicit in the acceptance of theories?
6. How compare different conceptions of truth? Are compositional truth theories always stronger than disquotational ones? How compare classical with the various nonclassical theories?
7. Is the theory of truth finitely axiomatizable?
8. Can a truth theory be categorical in some sense?
9. How are semantic concepts like compositionality related to mathematical concepts such as predicativity?

Tarskian disquotation

$\mathsf{TB}\restriction$ contains all axioms of PA plus all sentences

$$\mathsf{T}\ulcorner\varphi\urcorner \leftrightarrow \varphi$$

for all sentences φ without T. $\mathsf{TB}\restriction$ has only induction without T.

$\ulcorner\varphi\urcorner$ is the numeral of the code of the sentence φ.

This is the minimal theory that is adequate in the sense of Convention T (after adding $\forall x\,(\mathsf{T}x \to \mathrm{Sent}(x))$), i.e, 'Only T-free sentences are true'.

Theorem

$\mathsf{TB}\restriction$ is conservative over PA and thus consistent (Tarski). The truth predicate T isn't definable in PA (undefinability of truth). Any model of PA can be extended to a model of $\mathsf{TB}\restriction$.

'conservative' means that $\mathsf{TB}\restriction$ does not prove more arithmetical theorems than PA itself.

Tarskian disquotation with full induction

TB contains all axioms of PA *including all induction axioms with* T plus all sentences

$$\mathsf{T}\ulcorner\varphi\urcorner \leftrightarrow \varphi$$

for all sentences φ without T.

Theorem

TB is still conservative over PA. But not any model of PA can be extended to a model of TB (Engström 2009, but unpublished, Cieśliński 2015).

uniform Tarskian disquotation with full induction

UTB contains all axioms of PA including all induction axioms with T plus all sentences

$$\forall t_1 \dots \forall t_n \left(\mathsf{T}\ulcorner \varphi(\dot{t}_1, \dots, \dot{t}_n)\urcorner \leftrightarrow \varphi(t_1{}^\circ, \dots, t_n{}^\circ)\right)$$

where $\varphi(x_1, \dots, x_n)$ is a formula without T with exactly $x_1, \dots, x_n$ free.

The quantifier $\forall t$ ranges over all closed terms. t° is the value of t; there is no function symbol for $^\circ$, but we have a formula.

Theorem

UTB is conservative over PA, but not every model of PA can be expanded to a model of UTB.

Typed disquotational theories are weak (conservative over PA). Also, they don't prove generalizations such as:

$$\forall x\, \forall y\, \big(\mathrm{Sent}(x \dot{\wedge} y) \to (\mathsf{T}(x \dot{\wedge} y) \leftrightarrow \mathsf{T}(x) \wedge \mathsf{T}(y))\big)$$

Here $\mathrm{Sent}(x)$ expresses that x is a sentences of arithmetic (without T). $\dot{\wedge}$ expresses the function that yields, applied to formulæ φ and ψ their conjunction $(\varphi \wedge \psi)$.

Tarski (1935) observed this already and made some inconsistent claims. More of this later.

Untyped Disquotational Theories

Unlike their typed counterparts, untyped disquotational theories can be very strong.

Theorem

Any theory extending PA can reaxiomatized by a set of disquotation sentences $\mathsf{T}\ulcorner\varphi\urcorner \leftrightarrow \varphi$ over PA (McGee (1992) using a variant of Curry's paradox).

But these theories are not well motivated.

uniform Tarski disquotation for T-positive sentences

PUTB contains all axioms of PA including all induction axioms with T plus all sentences

$$\forall t_1 \dots \forall t_n \left(\mathsf{T}\ulcorner\varphi(\dot{t}_1, \dots, \dot{t}_n)\urcorner \leftrightarrow \varphi(t_1{}^\circ, \dots, t_n{}^\circ)\right)$$

$\varphi(x_1, \dots, x_n)$ is a formula of the language with T with exactly $x_1, \dots, x_n$ free such that T does not occur in the scope of $\neg$ ($\wedge$ and $\vee$ are the only connectives).

Theorem

PUTB is equivalent to $\mathrm{RA}_{<\epsilon_0}$ and KF below.

Theorem

PTB, in contrast, is conservative over PA (Cieśliński 2011).

There are reasonable, stronger disquotational theories. For instance, one can get second-order arithmetic (without second-order parameters) from a disquotational truth theory (Schindler 2015).

Generally, disquotational theories can vary significantly in their properties, depending on how the paradoxes are dealt with.

The reasonable disquotational theories mentioned so far fail to prove generalizations such as

$$\forall x\, \forall y \left(\mathrm{Sent}(x \mathbin{\underset{\cdot}{\land}} y) \rightarrow (\mathsf{T}(x \mathbin{\underset{\cdot}{\land}} y) \leftrightarrow \mathsf{T}(x) \land \mathsf{T}(y)) \right)$$

Compositionality via Reflection

Halbach (2001), Horsten and Leigh (2015), and others have tried to derive such compositional principles from disquotational theories by adding proof-theoretic reflection. Adding reflection is usual justified by appealing to implicit commitment (see Dean 2015, Cieśliński 2017, Fischer et al. 2021).

The reasonable disquotational theories mentioned so far fail to prove generalizations such as

$$\forall x\, \forall y\, \big(\mathrm{Sent}(x \dot{\land} y) \to (\mathsf{T}(x \dot{\land} y) \leftrightarrow \mathsf{T}(x) \land \mathsf{T}(y))\big)$$

Compositionality via Reflection

Halbach (2001), Horsten and Leigh (2015), and others have tried to derive such compositional principles from disquotational theories by adding proof-theoretic reflection. Adding reflection is usual justified by appealing to implicit commitment (see Dean 2015, Cieśliński 2017, Fischer et al. 2021).

Typed Compositional Theories

Other have tried to add compositional principles as axioms, although Tarski was dismissive of such attempts.

Compositional axioms may allow finite axiomatizability.

typed compositional truth

The system CT↾ is given by all the axioms of PA and the following axioms:

CT1 $\forall s\,\forall t\,\big(\mathsf{T}(s \underset{\cdot}{=} t) \leftrightarrow s^{\circ} = t^{\circ}\big)$ (= is the only predicate in PA)

CT2 $\forall x\,\big(\mathrm{Sent}(x) \rightarrow (\mathsf{T}(\underset{\cdot}{\neg} x) \leftrightarrow \neg \mathsf{T} x)\big)$

CT3 $\forall x\,\forall y\,\big(\mathrm{Sent}(x \underset{\cdot}{\wedge} y) \rightarrow (\mathsf{T}(x \underset{\cdot}{\wedge} y) \leftrightarrow \mathsf{T}(x) \wedge \mathsf{T}(y))\big)$

CT4 $\forall x\,\forall y\,\big(\mathrm{Sent}(x \underset{\cdot}{\vee} y) \rightarrow (\mathsf{T}(x \underset{\cdot}{\vee} y) \leftrightarrow \mathsf{T}(x) \vee \mathsf{T}(y))\big)$

CT5 $\forall v\,\forall x\,\big(\mathrm{Sent}(\underset{\cdot}{\forall} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\forall} v x) \leftrightarrow \forall t\,\mathsf{T}(x(t/v)))\big)$

CT6 $\forall v\,\forall x\,\big(\mathrm{Sent}(\underset{\cdot}{\exists} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\exists} v x) \leftrightarrow \exists t\,\mathsf{T}(x(t/v)))\big)$

Induction is restricted to sentences without T.

Theorem (Kotlarski et al. 1981, Enayat and Visser 2015, Leigh 2015)

CT↾ is conservative over PA.

Definition

A type over a model $\mathcal{M}$ is a finitely satisfied set of formulae $\varphi(x, \overline{b})$ that have exactly the variable x free and contain at most the parameter $\overline{b}$ for one fixed object $b \in |\mathcal{M}|$. A type p is recursive if and only if the set of codes of formulae $\varphi(x, y)$ with $\varphi(x, \overline{b}) \in p$ is recursive.

Definition (recursive saturation)

A model $\mathcal{M}$ of Peano arithmetic is recursively saturated if and only if every recursive type over $\mathcal{M}$ is realized (i.e. satisfied).

Theorem (Kotlarski et al. 1981, Lachlan 1981)

For all countable models $\mathcal{M}$ of PA:

$\mathcal{M}$ can be expanded to a model of $\mathsf{CT}\restriction$ iff $\mathcal{M}$ is recursively saturated.

If a model of PA can be expanded to a model of CT↾ at all (in which case it will be rec. saturated), there will be uncountably many ways to do so.

Also, CT↾ does not prove that all sentences of the form

$$0=0 \wedge 0=0 \wedge 0=0 \wedge \ldots \wedge 0=0$$

are true.

We also cannot prove that (the universal closures of) all theorems of PA are true, because we cannot prove that all arithmetical instances of induction are true and because we lack induction over the length of proofs.

So let's add induction.

If a model of PA can be expanded to a model of $\mathsf{CT}{\restriction}$ at all (in which case it will be rec. saturated), there will be uncountably many ways to do so.

Also, $\mathsf{CT}{\restriction}$ does not prove that all sentences of the form

$$0 = 0 \wedge 0 = 0 \wedge 0 = 0 \wedge \ldots \wedge 0 = 0$$

are true.

We also cannot prove that (the universal closures of) all theorems of PA are true, because we cannot prove that all arithmetical instances of induction are true and because we lack induction over the length of proofs.

So let's add induction.

If a model of PA can be expanded to a model of CT↾ at all (in which case it will be rec. saturated), there will be uncountably many ways to do so.

Also, CT↾ does not prove that all sentences of the form

$$0=0 \wedge 0=0 \wedge 0=0 \wedge \ldots \wedge 0=0$$

are true.

We also cannot prove that (the universal closures of) all theorems of PA are true, because we cannot prove that all arithmetical instances of induction are true and because we lack induction over the length of proofs.

So let's add induction.

typed compositional truth with full induction

CT is CT$\restriction$ plus all induction axioms in the language with T.

Using induction with T one proves in CT that all axioms of PA are true, and then using induction again, one proves that all *theorems* of PA are true. Since $0 \neq 1$ is not true, CT proves that PA is consistent.

Theorem

CT is no longer conservative over PA. The effect of adding the CT axioms to PA is the same as adding elementary comprehension. CT is equivalent to ACA.

The compositional axioms are required to handle parameters in comprehension. UTB interprets parameter-free elementary comprehension (and gives neat proof of the conservativity of ACA$\restriction$).

There are claims in the literature that the compositional axioms fix the extension of the truth predicate; but no truth theory containing TB does (Beth's theorem).

But CT 'fixes the extension of the truth predicate' in the following way:

Theorem

Let CT' be CT with the predicate T' instead of T. $\mathsf{CT} \cup \mathsf{CT}'$ *plus induction in the mixed language* proves $\forall x \left(\mathrm{Sent}(x) \to (\mathsf{T}x \leftrightarrow \mathsf{T}'x)\right)$.

To get even stronger theories one can iterate the theory CT along some ordinal notation system. The system turns out to be equivalent with iterated predicative comprehension (with some qualifications) (Feferman 1991, Halbach 2014, Fujimoto 2010).

Untyped Compositional Theories

Type-free compositional theories are usually extensions of CT, but there are also axioms about the truth of sentences with T.

Of course, here we need to heed the paradoxes. There is not only the liar paradox. Truth theories can be ω-inconsistent or have only trivial models.

All theories below have full induction.

Untyped Compositional Theories

Type-free compositional theories are usually extensions of CT, but there are also axioms about the truth of sentences with T.

Of course, here we need to heed the paradoxes. There is not only the liar paradox. Truth theories can be ω-inconsistent or have only trivial models.

All theories below have full induction.

Untyped Compositional Theories

Type-free compositional theories are usually extensions of CT, but there are also axioms about the truth of sentences with T.

Of course, here we need to heed the paradoxes. There is not only the liar paradox. Truth theories can be ω-inconsistent or have only trivial models.

All theories below have full induction.

FS is a system of *classical* and *symmetric* truth.

Friedman–Sheard

The system FS has all axioms of PA plus the following axioms and rules:

FS1 $\forall s\,\forall t\,\big(\mathsf{T}(s \mathbin{\underset{\cdot}{=}} t) \leftrightarrow s^\circ = t^\circ\big)$

FS2 $\forall x\,\big(\,\mathrm{Sent}_{\mathsf{T}}(x) \rightarrow (\mathsf{T}\underset{\cdot}{\neg}x \leftrightarrow \neg\mathsf{T}x)\big)$

FS3 $\forall x\,\forall y\,\big(\,\mathrm{Sent}_{\mathsf{T}}(x \mathbin{\underset{\cdot}{\wedge}} y) \rightarrow (\mathsf{T}(x \mathbin{\underset{\cdot}{\wedge}} y) \leftrightarrow (\mathsf{T}x \wedge \mathsf{T}y))\big)$

FS4 $\forall x\,\forall y\,\big(\,\mathrm{Sent}_{\mathsf{T}}(x \mathbin{\underset{\cdot}{\vee}} y) \rightarrow (\mathsf{T}(x \mathbin{\underset{\cdot}{\vee}} y) \leftrightarrow (\mathsf{T}x \vee \mathsf{T}y))\big)$

FS5 $\forall v\,\forall x\,\big(\,\mathrm{Sent}_{\mathsf{T}}(\underset{\cdot}{\forall} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\forall} v x) \leftrightarrow \forall t\,\mathsf{T}(x(t/v)))\big)$

FS6 $\forall v\,\forall x\,\big(\,\mathrm{Sent}_{\mathsf{T}}(\underset{\cdot}{\exists} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\exists} v x) \leftrightarrow \exists t\,\mathsf{T}(x(t/v)))\big)$

$$\text{NEC}\quad \frac{\varphi}{\mathsf{T}\ulcorner\varphi\urcorner} \qquad\qquad \frac{\mathsf{T}\ulcorner\varphi\urcorner}{\varphi}\quad \text{CONEC}$$

$\mathrm{Sent}_{\mathsf{T}}(x)$ expresses that x is a sentence possibly containing T.

Theorem

The liar sentence is neither provable nor refutable in FS.

CONEC is not needed for proof-theoretic strength; NEC can be expressed via iterated reflection.

Natural models of FS can be obtained via finitely iterated revision (in Gupta–Herzberger style).

FS is ω-inconsistent (McGee 1985)

FS is equivalent to finitely iterated Tarskian truth, i.e., iterated CT. Every application of NEC adds one layer of truth.

Kripke–Feferman

KF is PA with full induction plus the following axioms:

KF1 $\forall s\,\forall t\,\big(\mathsf{T}(s \underset{\cdot}{=} t) \leftrightarrow s^{\circ} = t^{\circ}\big)$

KF2 $\forall s\,\forall t\,\big(\mathsf{T}(\underset{\cdot}{\neg} s \underset{\cdot}{=} t) \leftrightarrow s^{\circ} \neq t^{\circ}\big)$

KF3 $\forall x\,\big(\mathrm{Sent}_{\mathsf{T}}(x) \rightarrow (\mathsf{T}(\underset{\cdot}{\neg}\underset{\cdot}{\neg} x) \leftrightarrow \mathsf{T}x)\big)$

KF4 $\forall x\,\forall y\,\big(\mathrm{Sent}_{\mathsf{T}}(x \underset{\cdot}{\wedge} y) \rightarrow (\mathsf{T}(x \underset{\cdot}{\wedge} y) \leftrightarrow \mathsf{T}x \wedge \mathsf{T}y)\big)$

KF5 $\forall x\,\forall y\,\big(\mathrm{Sent}_{\mathsf{T}}(x \underset{\cdot}{\wedge} y) \rightarrow (\mathsf{T}\underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y) \leftrightarrow \mathsf{T}\underset{\cdot}{\neg} x \vee \mathsf{T}\underset{\cdot}{\neg} y)\big)$

KF6 $\forall x\,\forall y\,\big(\mathrm{Sent}_{\mathsf{T}}(x \underset{\cdot}{\vee} y) \rightarrow (\mathsf{T}(x \underset{\cdot}{\vee} y) \leftrightarrow \mathsf{T}x \vee \mathsf{T}y)\big)$

KF7 $\forall x\,\forall y\,\big(\mathrm{Sent}_{\mathsf{T}}(x \underset{\cdot}{\vee} y) \rightarrow (\mathsf{T}\underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y) \leftrightarrow \mathsf{T}\underset{\cdot}{\neg} x \wedge \mathsf{T}\underset{\cdot}{\neg} y)\big)$

KF8 $\forall v\,\forall x\,\big(\mathrm{Sent}_{\mathsf{T}}(\underset{\cdot}{\forall} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\forall} v x) \leftrightarrow \forall t\,\mathsf{T}(x(t/v)))\big)$

KF9 $\forall v\,\forall x\,\big(\mathrm{Sent}_{\mathsf{T}}(\underset{\cdot}{\forall} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\neg}\underset{\cdot}{\forall} v x) \leftrightarrow \exists t\,\mathsf{T}(\underset{\cdot}{\neg} x(t/v)))\big)$

KF10 $\forall v\,\forall x\,\big(\mathrm{Sent}_{\mathsf{T}}(\underset{\cdot}{\exists} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\exists} v x) \leftrightarrow \exists t\,\mathsf{T}(x(t/v)))\big)$

KF11 $\forall v\,\forall x\,\big(\mathrm{Sent}_{\mathsf{T}}(\underset{\cdot}{\exists} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\neg}\underset{\cdot}{\exists} v x) \leftrightarrow \forall t\,\mathsf{T}(\underset{\cdot}{\neg} x(t/v)))\big)$

KF12 $\forall t\,(\mathsf{T}(\underset{\cdot}{\mathsf{T}} t) \leftrightarrow \mathsf{T}t^{\circ})$

KF13 $\forall t\,\big(\mathsf{T}\underset{\cdot}{\neg}\underset{\cdot}{\mathsf{T}} t \leftrightarrow (\mathsf{T}\underset{\cdot}{\neg} t^{\circ} \vee \neg\,\mathrm{Sent}_{\mathsf{T}}(t^{\circ}))\big)$

KF14 $\forall x\,\big(\mathrm{Sent}_{\mathsf{T}}(x) \rightarrow \neg(\mathsf{T}x \wedge \mathsf{T}\underset{\cdot}{\neg} x)\big)$

The last axiom rules out truth-value gluts.

KF axiomatizes Kripke's (1975) theory of truth with Strong Kleene.

Several variants of KF can be found in the literature. This is my version.

The last axiom – called the consistency axiom – excludes truth-value gluts.

Theorem

KF is an axiomatization of Kripke's theory of truth with the Strong Kleene schema.

It's an axiomatization of a partial notion of truth in classical logic.

KF proves the liar sentence, as it excludes truth-value gluts.

Even without the consistency axiom KF14, KF cannot be closed consistently under NEC and CONEC. It's essentially asymmetric.

KF is equivalent to $\epsilon_0 = \omega^{\omega^{\cdot^{\cdot^{\omega}}}}$ iterated Tarskian truth (Feferman 1991)) Feferman's analysis proceeds in terms of infinite conjunctions.

Further systems:

- Variations of KF: Feferman's (1991) strong reflective closure of PA, weak Kleene, Feferman's (2008) DT
- Cantini's (1990) VF and supervaluations

Further systems:

- Variations of KF: Feferman's (1991) strong reflective closure of PA, weak Kleene, Feferman's (2008) DT
- Cantini's (1990) VF and supervaluations

CD is PA with full induction plus the following axioms:

(T1) $\forall s\, \forall t\, \big(\mathsf{T} s \underset{\cdot}{=} t \leftrightarrow s^\circ = t^\circ\big)$

(T2) $\forall t\, (\mathsf{T}\underset{\cdot}{\mathsf{D}} t \leftrightarrow \mathsf{D} t^\circ)$

(T3) $\forall t\, \big(\mathsf{D} t^\circ \rightarrow (\mathsf{T}\underset{\cdot}{\mathsf{T}} t \leftrightarrow \mathsf{T} t^\circ)\big)$

(T4) $\forall x\, \big(\mathrm{Sent}(x) \rightarrow (\mathsf{T}(\underset{\cdot}{\neg} x) \leftrightarrow \neg \mathsf{T} x)\big)$

(T5) $\forall x\, \forall y\, \big(\mathrm{Sent}(x \underset{\cdot}{\wedge} y) \rightarrow (\mathsf{T}(x \underset{\cdot}{\wedge} y) \leftrightarrow \mathsf{T} x \wedge \mathsf{T} y)\big)$

(T6) $\forall v\, \forall x\, \big(\mathrm{Sent}(\underset{\cdot}{\forall} v x) \rightarrow (\mathsf{T}(\underset{\cdot}{\forall} v x) \leftrightarrow \forall t\, \mathsf{T} x[t/v])\big)$

(D1) $\forall s\, \forall t\, \mathsf{D}(s \underset{\cdot}{=} t)$

(D2) $\forall t\, (\mathsf{D}\underset{\cdot}{\mathsf{D}} t \leftrightarrow \mathsf{D} t^\circ)$

(D3) $\forall s\, \big(\mathsf{D}\underset{\cdot}{\mathsf{T}} s \leftrightarrow \mathsf{D} s^\circ\big)$

(D4) $\forall x\, \big(\mathrm{Sent}(x) \rightarrow (\mathsf{D}(\underset{\cdot}{\neg} x) \leftrightarrow \mathsf{D} x)\big)$

(D5) $\forall x\, \forall y\, \Big(\mathrm{Sent}(x \underset{\cdot}{\wedge} y) \rightarrow \Big(\mathsf{D}(x \underset{\cdot}{\wedge} y) \leftrightarrow \big((\mathsf{D} x \wedge \mathsf{D} y) \vee (\mathsf{D} x \wedge \neg \mathsf{T} x) \vee (\mathsf{D} y \wedge \neg \mathsf{T} y)\big)\Big)\Big)$

(D6) $\forall v\, \forall x\, \Big(\mathrm{Sent}(\underset{\cdot}{\forall} v x) \rightarrow \Big((\mathsf{D}(\underset{\cdot}{\forall} v x) \leftrightarrow$
$\big(\forall t\, \mathsf{D} x[t/v] \vee \exists t (\mathsf{D} x[t/v] \wedge \neg \mathsf{T} x[t/v])\big)\Big)\Big)$

Typed disquotational theories are weak, but often have interesting model-theoretic properties. They are not as harmless as some deflationists and disquotationalists think.

Untyped disquotational theories can be very strong and, e.g., handle inductive definitions.

Typed compositional theory is related to arithmetical comprehension, but truth and comprehension are different, as $\mathsf{CT}\restriction$ shows. Systems with arithmetical induction only have interesting model-theoretic (rec. saturation) and proof-theoretic properties.

Untyped compositional theories can be strong and match the strength of impredicative theories such as ID_1; others are conservative over PA.

The paradoxes are not only a nuisance, they can be put to work to produce interesting results – just like elsewhere: Diagonalization can lead to disaster, but also very productive.

Which theory should philosophers pick?

For philosophical applications, the theory should be untyped and be fully classical (so not KF). Everything else requires an even more profound rewriting of philosophy.

Now for dissenting voices…

Which theory should philosophers pick?

For philosophical applications, the theory should be untyped and be fully classical (so not KF). Everything else requires an even more profound rewriting of philosophy.

Now for dissenting voices…

Andrea Cantini. A Theory of Formal Truth Arithmetically Equivalent to ID_1. *Journal of Symbolic Logic*, 55:244–259, 1990.

Cezary Cieśliński. T-equivalences for positive sentences. *The Review of Symbolic Logic*, 4:319–325, 2011.

Cezary Cieśliński. Typed and untyped disquotational truth. In Theodora Achourioti, Henri Galinon, José Martínez Fernández, and Kentaro Fujimoto, editors, *Unifying the Philosophy of Truth*, pages 307–320. Springer Netherlands, Dordrecht, 2015.

Cezary Cieśliński. *The Epistemic Lightness of Truth: Deflationism and its Logic*. Cambridge University Press, 2017.

Walter Dean. Arithmetical reflection and the provability of soundness. *Philosophia Mathematica*, 23:31–64, 2015.

Ali Enayat and Albert Visser. New Constructions of Satisfaction Classes. In T. Achourioti, H. Galinon, K. Fujimoto, and J. Martinez-Fernandez, editors, *Unifying the Philosophy of Truth*, pages 321–335. Springer, 2015.

Solomon Feferman. Towards Useful Type-Free Theories I. *Journal of Symbolic Logic*, 49:75–111, 1984.

Solomon Feferman. Reflecting on Incompleteness. *Journal of Symbolic Logic*, 56:1–49, 1991.

Solomon Feferman. Axioms for determinateness and truth. *Review of Symbolic Logic*, 1:204–217, 2008.

Hartry Field. *Saving Truth From Paradox*. Oxford University Press, Oxford, 2008.

Martin Fischer, Leon Horsten, and Carlo Nicolai. Hypatia's silence. *Noûs*, 55:62–85, 2021.

Kentaro Fujimoto. Relative truth Definability. *Bulletin of Symbolic Logic*, 16:305–344, 2010.

Volker Halbach. Disquotational Truth and Analyticity. *Journal of Symbolic Logic*, 66:1959–1973, 2001.

Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, revised edition, 2014. first edition 2011.

Volker Halbach and Leon Horsten. Axiomatizing Kripke's Theory of Truth. *Journal of Symbolic Logic*, 71:677–712, 2006.

Volker Halbach and Graham Leigh. *The Road to Paradox: A Guide to Syntax, Truth, and Modality*. Cambridge University Press, 2022. to appear.

Leon Horsten and Graham E. Leigh. Truth is Simple, 2015.

Henryk Kotlarski, Stanislav Krajewski, and Alistair Lachlan. Construction of Satisfaction Classes for Nonstandard Models. *Canadian Mathematical Bulletin*, 24:283–293, 1981.

Michael Kremer. Kripke and the logic of truth. *Journal of Philosophical Logic*, 17:225–278, 1988.

Saul Kripke. Outline of a Theory of Truth. *Journal of Philosophy*, 72:690–716, 1975. reprinted in Martin (1984).

Alistair Lachlan. Full Satisfaction Classes and Recursive Saturation. *Canadian Mathematical Bulletin*, 24:295–297, 1981.

Graham Leigh and Michael Rathjen. The Friedman–Sheard programme in intuitionistic logic. *Journal of Symbolic Logic*, 77:777–806, 2012.

Graham E. Leigh. Conservativity for theories of compositional truth via cut elimination. *jsl*, 80:825–865, 2015.

Robert L. Martin, editor. *Recent Essays on Truth and the Liar Paradox*. Clarendon Press and Oxford University Press, Oxford and New York, 1984.

Vann McGee. How Truthlike Can a Predicate Be? A Negative Result. *Journal of Philosophical Logic*, 14:399–410, 1985.

Vann McGee. Maximal consistent sets of instances of Tarski's schema (T). *Journal of Philosophical Logic*, 21:235–241, 1992.

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

Alfred Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica Commentarii Societatis Philosophicae Polonorum*, 1:261–405, 1935. translated as 'The Concept of Truth in Formalized Languages' in (Tarski 1956, 152–278); page references are given for the translation.

Alfred Tarski. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. Clarendon Press, Oxford, 1956.