Truth in Metamathematics

Leon Horsten
University of Bristol

Amsterdam Workshop on Truth
13–15 March 2013
Overview

[This talk is based on joint work with Martin Fischer]

1. The purpose of truth theories
2. Expressiveness
3. Conservativeness
4. Minimally adequate truth theories
5. Finite axiomatisability
6. Completeness
7. Speed-up
8. Reflection principles
9. References
The purpose of truth theories

Question

*What is the best (a good) theory of truth?*
The purpose of truth theories

Question

What is the best (a good) theory of truth?

Depends on the purpose . . .

- truth in natural language
- truth in philosophy
- truth in mathematics
- . . .
Tarski on truth in natural language

- material adequacy:
  
  the Tarski biconditionals

- natural language is semantically closed

- $\Rightarrow$ natural language is inconsistent
Tarski on formalising metamathematics

- Truth is used in metamathematics (model theory)
- Metamathematics is coherent
  - Tarski thought that this needs to be proved

⇒ Explicate the role of truth in metamathematics
The expressive role of truth in metamathematics

Thesis

The concept of truth plays an expressive role in metamathematics.

- the concept of expressive power is not precise
  - I will not try to give a precise definition of expressiveness.
Interpretability

- a concept that can be simulated in $T$ does not add real expressive power to $T$
- if $T$ is more expressive than $S$ then $T$ is non-interpretable in $S$
Interpretability

- a concept that can be simulated in $T$ does not add real expressive power to $T$
- if $T$ is more expressive than $S$ then $T$ is non-interpretable in $S$

Thesis (Condition 1)

A truth theory for metamathematics must be non-interpretable.
 Truth and higher types

► “The expressive role of truth can also be played by higher types.”
  ► Tarski’s definition of truth

► Objections:
  ► this is overkill
  ► what about set theory?
    ► class theory? …

⇒ what is minimally needed for doing metamathematics?
The algebraic stance

The semantical turn:

\[
\text{theory} = \text{class of models}
\]

Thesis

Model theory is an algebraic discipline: no models should be excluded.
Two notions of conservativeness

Definition
$T$ is a (proof theoretically) conservative extension of $S$ if and only if $S \subseteq T$ and for all $\varphi$ in the language of $S$, if $T$ proves $\varphi$, then already $S$ proves $\varphi$.

Definition
$T$ is a (semantically) conservative extension of $S$ if and only if all the models of $S$ can be expanded to models of $T$. 
Two notions of conservativeness

Definition

$T$ is a (proof theoretically) conservative extension of $S$ if and only if $S \subseteq T$ and for all $\varphi$ in the language of $S$, if $T$ proves $\varphi$, then already $S$ proves $\varphi$.

Definition

$T$ is a (semantically) conservative extension of $S$ if and only if all the models of $S$ can be expanded to models of $T$.

Thesis (Condition 2)

A truth theory for metamathematics must be semantically conservative
Proof theoretic and semantic conservativeness are not equivalent:

- semantic conservativeness $\Rightarrow$ proof theoretic conservativeness
- proof theoretic conservativeness $\not\Rightarrow$ semantic conservativeness

$\text{CT} \upharpoonright$: the Tarskian compositional theory of truth with induction restricted to the truth-free background language of arithmetic

- $\text{CT} \upharpoonright$ is proof theoretically conservative over $PA$ but not semantically conservative over $PA$

$\Rightarrow$ $\text{CT} \upharpoonright$ is not a satisfactory truth theory for metamathematics
Pulling in opposite directions

Question

*Which theory of truth is a suitable framework for doing metamathematics?*

- non-interpretable: strong truth theory
- semantically conservative: weak truth theory

Can a theory be strong and weak (in the relevant senses) at the same time?
The theory $\text{PA}^-$

- typed: truth restricted to arithmetical formulae
- compositional: e.g.,

$$\forall \phi, \psi \in \mathcal{L}_{\text{PA}} : T(\phi \land \psi) \leftrightarrow (T\phi \land T\psi)$$

- positive: negation does not commute with truth
- induction restricted to total formulas:

$$\text{tot}(\phi) \equiv T\phi \lor T\neg\phi$$

**Thesis**

$PT^- \text{ is a suitable truth theory for metamathematics.}$
Properties

Proposition

\( PT^- \) is interdefinable with \( ACA_0 \)

- semantically conservative
- noninterpretable

In a (admittedly vague) sense \( PT^- \) is the truth theory corresponding to \( ACA_0 \)
Counterparts

- if we go type-free then Cantini’s theory $KF_t$ results
- the analogue for set theory ($KF_{tc}$) was investigated by Fujimoto
  - $KF_{tc}$ is the truth theory corresponding to $NBG$
Theorem

Any theory can be finitely axiomatised in a language expansion with one new predicate.
The proof of Craig-Vaught

- the new predicate is a truth predicate
- the proof is a model expansion argument
- $PT^-$ is a finite axiomatisation of its background theory
The strength of the completeness theorem

- The completeness theorem for deductively closed axiomatic theories needs $RCA_0$
- The completeness theorem for axiomatic theories needs $WKL_0$
- The textbook proof of the completeness theorem needs $ACA_0$
Completeness and $PT^-$

- $WKL_0$ is interpretable in $PA$
- $\Rightarrow$ the proof of completeness does not require a non-interpretable notion of truth
Completeness and $PT^-$

- $WKL_0$ is interpretable in $PA$
- $\implies$ the proof of completeness does not require a non-interpretable notion of truth

However:

**Proposition**

*The textbook proof of completeness can be expressed (in a natural way) in $PT^-$, but not in a weaker truth theory.*
Speed up and expressive power

Having non-trivial speed up is an indication of expressive power
Proposition

$PT^−$ has non-trivial speed up over $PA$.

Whether $CT|^\uparrow$ has nontrivial speed up is an open problem
Reflection on a cut

$PT^-$ cannot prove full reflection but only a restricted version of it.

**Proposition**

$PT^-$ can prove reflection for PA on a cut.
References