# The innocence of truth

Cezary Cieśliński

Institute of Philosophy
University of Warsaw
Poland

Helsinki 2015

## The topic

**Deflationary claim:** the notion of truth is innocent or “metaphysically thin”.

**One explication**: an adequate theory of truth for a given language conservatively extends a base theory of syntax for this language (papers by L. Horsten, S. Shapiro, and J. Ketland).

**Basic observations:**

- Conservativity claims were put forward not by the deflationists but by their critics.
- Since then conservativity requirement has taken a life of its own.

**Main question:** in what sense – if any – do conservativity claims form a part of deflationist doctrines?

## Two notions of conservativity

Let $T_1$ and $T_2$ be theories in languages $L_1$ and $L_2$ (with $L_1 \subseteq L_2$).

### Syntactic conservativity

$T_2$ is syntactically conservative over $T_1$ iff $T_1 \subseteq T_2$ and $\forall \psi \in L_1 [T_2 \vdash \psi \rightarrow T_1 \vdash \psi]$.

### Semantic conservativity

$T_2$ is semantically conservative over $T_1$ iff every model of $T_1$ can be expanded to a model of $T_2$.

**Remark:** if $T_2$ is semantically conservative over $T_1$, then $T_2$ is syntactically conservative over $T_1$. The opposite implication does not hold!

## Conservativity requirement: the basic intuitions

*How thin can the notion of arithmetic truth be, if by invoking it we can learn more about the natural numbers?*

*Suppose, for example, that Karl [...] adds a truth predicate to the language and extends [a base theory]* $B$ *to a [truth] theory* $B'$ *[...] Assume that* $B'$ *is not conservative over* $B$*. Then there is a sentence* $\Phi$ *in the original language (so that* $\Phi$ *does not contain the truth predicate) such that* $\Phi$ *is a consequence of* $B'$ *but not a consequence of* $B$*. That is, it is logically possible for the axioms of* $B$ *to be true and yet* $\Phi$ *false, but it is not logically possible for the axioms of* $B'$ *to be true and* $\Phi$ *false. This undermines the central deflationist theme that truth is in-substantial.*

**Source:** S. Shapiro, "Proof and truth: through thick and thin", *Journal of Philosophy* 1998, 95(10), 493-521.

## Semantic conservativity: intuitions

*If an extension of the domain or some structural change of it were required, then truth would exhibit extralinguistic effects. It would affect the things the language talks about and not just our way of speaking of them. [...]*

*Not only would we operate at a linguistic level by adding a new expression and interpreting it with a suitable extension, we would also need to intervene into the domain by changing and shaping it. In this sense, and in open contrast with the deflationist claim, the property of truth would enter reality as a robust ingredient.*

**Source:** A. Strollo, "Deflationism and the invisible power of truth", *Dialectica* 2013, 67, 521-543.

## Conservative extensions: basic examples.

**Notation**

- $L_{PA}$ is the language of Peano arithmetic.
- $L_T$ is the result of extending $L_{PA}$ with a new predicate "$T$".
- $Ind_{\psi(x)}$ is the axiom of induction for $\psi(x)$.

**Definition**

$TB = PA \cup \{T(\varphi) \equiv \varphi : \varphi \in L_{PA}\} \cup \{Ind_{\psi(x)} : \psi(x) \in L_T\}$.

**Fact**

**(a)** *TB is syntactically conservative over PA.*

**(b)** *TB is not semantically conservative over PA.*

## Syntactic conservativity of *TB*

A very simple argument consists in observing, that in a given proof of an arithmetical sentence in *TB*, all occurrences of "*T*" can be replaced by a suitable arithmetically definable, partial truth predicate.

**Remark:** for TB *without* extended induction, it is possible to prove semantic conservativity as well. For an arbitrary model *M*, define:

$$S = \{\psi \in L_{PA} : M \models \psi\}.$$

Then $(M, S)$ is a model for all T-biconditionals in *TB*. However, for a nonstandard *M*, the set *S* is not inductive.

## Semantic non-conservativity of *TB*

The following two conditions are equivalent for an arbitrary nonstandard model $M$ of Peano arithmetic:

***(i)*** there is a set $S \subseteq M$ such that $(M, S) \models TB$;

***(ii)*** $M$ codes $Th(M)$, i.e. the set of all sentences true in $M$ is coded by an element of $M$.

In view of the equivalence between *(i)* and *(ii)*, we obtain:

> *TB is semantically conservative over PA iff every nonstandard model of PA codes its own theory.*

Since models of *PA* which do not code their theories are known to exist, there are models of *PA* which cannot be expanded to models of *TB*.
In other words: *TB* is not semantically conservative over *PA*.

## Another example - theory *CT*

### Definition

Apart from the axioms of *PA* and induction axioms for the extended language, *CT* contains the following truth axioms:

- $\forall s, t \in Tm^c \big(T(s = t) \equiv val(s) = val(t)\big)$
- $\forall x, \big(Sent_{L_{PA}}(x) \to (T\neg x \equiv \neg Tx)\big)$
- $\forall x, \forall y, \big(Sent_{L_{PA}}(x \wedge y) \to (T(x \wedge y) \equiv (Tx \wedge Ty))\big)$
- $\forall v, \forall x, \big(Sent_{L_{PA}}(\forall vx) \to (T(\forall vx) \equiv \forall t T(x(t/v)))\big)$

A theory like *CT* but with arithmetical induction only will be denoted as $CT^-$.

### Fact

*$CT^-$ is syntactically, but not semantically conservative over PA. CT is not syntactically conservative over PA (it proves e.g. "$Con_{PA}$").*

## Conservativity: general remarks

1. Non-conservativity phenomena are not associated solely with compositional truth theories.
2. Fully inductive truth theories are never innocent.
3. For disquotational truth theories (*TB*, *UTB*, *PTB*, *PUTB*), removing extended induction produces typically a semantically conservative extension. Compositional theories are often different in this respect.
4. Semantic conservativity can be squared with compositionality and *some* extended induction (example: theory $PT^-$ with positive compositional axioms and a restricted form of extended induction).

## Why should models matter?

Why should we demand admissibility of *all* models?

1. Arithmetical truth is truth in some model (or a class of models) of Peano Arithmetic, corresponding to a fragment of real world.
2. We have no way of recognizing models which do not correspond to a fragment of the real world.
3. A theory which excludes some models risks excluding the model corresponding to the real world.
4. Therefore, all models should be treated on a par.

**Problem:** Arithmetical truth is presented in premise 1 as truth in some special (intended) models. In effect, a stronger notion of truth than the one characterized by the deflationary axioms is used here to justify the conservativity demand. However, the deflationists claim that stronger notions of truth are not needed.

## Why should models matter, continued

**Alternative approach:** we could declare that the notion of the intended model is incomprehensible and that all models are on a par.

**Comments:**

- Such a move is natural in some contexts (e.g. first order logic, group theory);
- It's very risky in other contexts. For example, there is the intuition that models of $PA + \neg Con_{PA}$ get the arithmetic *wrong*. (No analogy with the result of adding "the operation is commutative" to group axioms).

## Why syntactic conservativity?

Shapiro and Ketland: the deflationists claim that truth is just a tool which in principle can be disposed of in explanations or justifications of non-semantic facts.

Some textual support:

> *On this issue, contemporary deflationists are in broad agreement: the function of truth talk is wholly expressive, thus never explanatory. As a device for semantic assent, the truth predicate allows us to endorse or reject sentences (or propositions) that we cannot simply assert, adding significantly to the expressive resources of our language. Of course, proponents of traditional theories of truth do not deny any of this. What makes deflationary views deflationary is their insistence that the importance of truth talk is exhausted by its expressive function.*

**Source:** M. Williams, “Meaning and deflationary truth”, *Journal of Philosophy* 1999, 96(11).

# Syntactic conservativity: the argument

1. Truth is never explanatory/justificatory.
2. If a theory of truth proves new non-semantic facts, then these new facts are explained/justified by truth-theoretic considerations.
3. Therefore a theory of truth does not prove new non-semantic facts, i.e. it is syntactically conservative over its base.

## Problems with explanation

> *A proof or proof sketch can give cogent grounds for believing a claim, but it might fail nonetheless to provide the sort of illumination we can hope for in mathematical investigation.*

**Source:** J. Tappenden, “Proof style and understanding in mathematics”, in *Visualization, Explanation and Reasoning Styles in Mathematics* Dordrecht, Springer, 2005.

**Main problems:**

- The concept of explanation in mathematics is at present neither well understood nor sufficiently studied.
- Is *every* mathematical proof explanatory? This is far from obvious. A separate argument would be needed showing that proofs employing truth axioms have this character.

## Problems with explanation

Consider a proof of a trivial theorem, e.g. $\varphi \to \varphi$ for arithmetical $\varphi$, carried out in $CT$. Idea of the proof:

- by compositional principles, derive "$\forall\psi T(\psi \to \psi)$",
- conclude that $T(\varphi \to \varphi)$,
- by disquotation (valid in $CT$) derive $\varphi \to \varphi$.

**Question:** is this proof explanatory?
(The proposed explanation: "we accept $\varphi \to \varphi$ because it's true, and it's true because truth commutes with sentential connectives.")

- If the answer is "yes", truth can have explanatory role even in conservative truth theories.
- If the answer is "no", why should e.g. proof of $Con_{PA}$ in $CT$ fare any better?

## Justificatory value of truth-theoretic arguments

**Main problem:** It might happen, that proofs of new non-semantic facts in a non-conservative theory of truth are not justificatory

**Example to consider:** a non-conservative theory $CT$, with full induction, proves the consistency of Peano arithmetic. How compelling is such a proof?

**Remark:** The question is not about the formal correctness of the proof of $Con_{PA}$ in $CT$ (the proof is correct!)

Imagine that someone has serious doubts about the consistency of $PA$. After seeing and understanding the proof of $Con_{PA}$ in $CT$, will he lose these doubts? Or, more importantly, *should* he lose them?

## Truth of axioms of *PA*

### Observation

*Let $Th = I\Sigma_1$ + compositional truth axioms + induction for $\Delta_0$ formulas of the extended language (with the truth predicate). Then Th proves: "all the axioms of PA are true".*

### Proof.

For the truth of inductive axioms, working in *Th* fix an arithmetical formula $\varphi(x)$ with one free variable. It is enough to obtain:

$$(*) \quad T(\varphi(0)) \wedge \forall x[T(\varphi(x)) \rightarrow T(\varphi(x+1))] \rightarrow \forall x T(\varphi(x)).$$

Then the truth of inductive axiom for $\varphi(x)$ follows by compositionality. Assume the antecedent of (*). For an indirect proof, assume also $\exists x \neg T(\varphi(x))$ and choose (using $\Delta_0$ induction) the smallest $x$ with this property. By the antecedent of (*) such a smallest $x$ can be neither zero nor a successor number, which generates a contradiction. □

## Consistency of *PA*

**Corollary**

*$I\Sigma_1$ + compositional truth axioms + induction for $\Pi_1$ formulas of the extended language proves: "all theorems of PA are true". Therefore this theory proves also the consistency of PA.*

**Comment:** $\Pi_1$ induction permits us to prove that all logical axioms are true (for arithmetical axioms, $\Delta_0$ induction is enough). It is unknown at the moment whether theory *Th* - with $\Delta_0$ induction only - proves global reflection for *PA*.

## Consistency justified?

### Observation

- *The proof requires some theory of syntax. However, it doesn't have to be full PA (the consistency of which is doubted here after all!).*
- *It is $\Delta_0$ (extended) induction that licenses a move from $\exists x \neg T(\varphi(x))$ to the choice of the smallest $x$ with this property.*
- *It does not matter that the extended induction is "just" $\Delta_0$: the principle works for an arbitrary arithmetical formula $\varphi(x)$, which is turned into a $\Delta_0$ formula by a mere quirk of syntax (i.e. by appending "$T$").*
- *Accepting the least number principle in this form is nothing short of accepting full arithmetical induction as credible.*

***Problem:*** *Someone who doubts the consistency of PA shouldn't be satisfied with a proof taking for granted full arithmetical induction as credible. Justificatory value of this proof is close to null.*

## Final observations

- Neither semantic nor syntactic conservativeness fares well as an explication of the traditional deflationary claims.
- In spite of this, conservativity is a *convenient* property.
- Syntactic conservativity can still function as a *new* explication of thinness, proposed with full awareness that its connection with the tradition is rather loose.

## References


- Cezary Cieśliński ‘The innocence of truth’, *dialectica* 69(1), 61–85, 2015.
- Cezary Cieśliński *The Epistemic Lightness of Truth. Deflationism and its Logic*, Cambridge University Press, 2017.

THE END

Thanks for your attention!!!