AXIOMATIZING TRUTH: HOW AND WHY

Solomon Feferman

Pillars of Truth Conference
Princeton University, April 8-10, 2011

Why Axiomatize?

1. Logical theories of truth comprise semantical (definitional) and axiomatic theories.
2. Various philosophical and semantical theories are candidates for axiomatization (but not all, e.g. coherence, pragmatic, fuzzy theories). NB: axiomatizations are not uniquely determined.
3. Only a few properties (e.g. T-scheme, compositionality) are uppermost in one's mind in spelling out phil. and sem. theories. Axiomatizing them brings out the consequent properties in full.

Why Axiomatize? (cont'd)

4. Axiomatic theories separate out the properties of a semantical construction from what is needed to justify that construction (e.g., set theory).
5. An axiomatization, if not of a sem. construction, can be proved consistent by providing a model.
6. Axiomatizations of phil. or sem. theories provide a framework within which to reason systematically about their properties, and thus assess them.

Why Axiomatize ? (final)

7. One can compare like and unlike axiomatizations as to their proof-theoretical strength, using an extensive body of well-established metamathematical techniques.

8. Given axiomatizations may be varied in natural ways, for example by extending general principles from one's base theory (e.g. induction in arithmetic, or separation in set theory, to the theory with a truth predicate.

How to Axiomatize: Leitgeb Criteria
"What theories of truth should be like (*but cannot be*)" (2007)

(L1) Truth should be expressible by a predicate (and a theory of syntax should be available).

(L2) If a theory of truth is added to mathematical or empirical theories, it should be possible to prove them true.

(L3) The truth predicate should not be subject to type restrictions.

(L4) T-biconditionals should be derivable unrestrictedly.

How Axiomatize? Leitgeb Criteria (cont'd)

(L5) Truth should be compositional.

(L6) The theory should allow for standard interpretations.

(L7) The outer logic and the inner logic should coincide.

(L8) The outer logic should be classical.

(L1)-(L8) Construed Axiomatically

- An axiomatic theory of truth is given by a formal system S.
- To meet (L1), assume S includes PA and L(S) contains the unary predicate symbol T(x). For A in L(S), #A = numeral of Gödel nr. of A. Write T(A) for T(#A).
- A minimum requirement for (L2) is that S proves P: "all sentences provable in PA are true".
- For (L3)(type free), allow all A from L(S) in T(A). Then in (L4), have all T(A) $\leftrightarrow$ A for such A.

(L1)-(L8) Construed Axiomatically (cont'd)

- Compositionality (L5) means that S proves T(¬A)↔ ¬T(A), and so on for the other connectives and quantifiers, possibly in variable form.
- (L6) means that S has a model expanding the standard model of PA.
- For (L7), the "outer logic" of S = the basic logical axioms and rules of S, and the "inner logic" = the laws holding of all A such that S proves T(A).
- In (L8), the outer logic is just classical logic.

What to Accept, What to Reject?

- Tarski's Undefinability Theorem If S satisfies (L1), (L4) and (L8), then S is inconsistent. (In fact, intuitionistic logic suffices).
- I accept (L1) (extension of PA), (L2) (prove all theorems of PA are true), (L3) (type free) and (L6) (has a model standard for PA).
- I accept (L4) (T-scheme) and (L5) (compositionality) *only in certain restricted forms*.
- I *reject* (L7) (outer logic = inner logic) and accept (L8) (classical outer logic).

How to Axiomatize? My Criteria
"Axioms for determinateness and truth"
Rev. Symbolic Logic (2008), cf. 206-207

- <u>The argument</u>: Every predicate has a domain D of significance; it makes sense to apply the predicate only to arguments of that domain.
- The domain D of significance of the truth predicate T consists of all propositions (via sentences) that are meaningful and determinate, i.e. have a definite truth value, true or false.
- D includes various sentences that involve the notion of truth, but not necessarily all.
- Note $T(A) \rightarrow D(A)$, all A; also $F(A) \rightarrow D(A)$ for all A, where $F(A)=T(\neg A)$.

How to Axiomatize? My criteria (cont'd)

(F1)-(F3) = (L1)-(L3), accepted.

(F4) The T-scheme is taken only in the following restricted form: $D(A) \rightarrow (T(A) \leftrightarrow A)$, for all A.

(F5) Similarly, compositionality holds only under the assumption of D for all formulas involved.

(F6) = (L6), accepted, but (L7) (outer logic = inner logic) is rejected. (F8) = (L8), accepted.

(F9) Though D(A) can be defined as $(T(A) \vee F(A))$, the conditions on D should be prior to (independent of) those on T.

Illustrations from three of my papers

I. The system DT (2008)

- DT is an extension of PA and meets all the criteria (F1)-(F6) and (F8). Its basic logical operations are $\neg$, $\vee$, $\rightarrow$ and $\forall$. Though the logic is classical, $\rightarrow$ is not defined in terms of $\neg$ and $\vee$.
- As required by (F9), the D axioms are prior to the T axioms for $\neg$, $\vee$, and $\forall$. For example we have $D(x\dot{\vee} y) \leftrightarrow D(x) \wedge D(y)$, and then

  $D(x\dot{\vee} y) \rightarrow [T(x\dot{\vee} y) \leftrightarrow T(x) \vee T(y)]$.

I. The System DT (cont'd)

- But for $\rightarrow$, we don't meet (F9) in full, only $D(x\rightarrow.y)\leftrightarrow D(x)\wedge[T(x)\rightarrow D(y)]$, though

  $D(x\rightarrow.y)\rightarrow[T(x\rightarrow.y)\leftrightarrow(T(x)\rightarrow T(y))]$ is standard.

- This has to do with (F2). Let P be the sentence $\forall x[\text{Prov-PA}(x)\wedge \text{Sent}(x)\rightarrow T(x)]$. If $\rightarrow$ were defined as usual in terms of $\neg$ and $\vee$, we can prove P. But the D condition above seems to be needed to prove T(P).

II. The System KF
“Reflecting on incompleteness”, *JSL* (1991)

- KF is an axiomatization of the Kripke 1975 definition of truth in Kleene 3-valued semantics that I made in 1979 and circulated then in notes. It was then studied by Reinhardt (1985), Cantini (1989) and McGee(1991) prior to my (1991).
- Its purpose was instrumental, to define a notion of *reflective closure* of a system S. That has more recently been superseded by a notion of the *unfolding of* S, without use of a theory of truth.
- But KF took on a life of its own as a theory of truth.

II. The System KF (cont'd

- KF violates criterion (L7) that requires the outer logic to = the inner logic, since the outer logic is classical while the inner logic is Kleene's 3-valued.
- For, with λ taken to be a "Liar" sentence, KF proves both $\lambda \vee \neg\lambda$ and $\neg T(\lambda \vee \neg\lambda)$.
- Why isn't this a problem, contrary to Leitgeb, Halbach, Horsten and others?

II. The System KF (1991) (cont'd)

(i) The distinction between outer and inner logics is a problem only if one conflates two notions of truth, namely Kripke's notion of *grounded truth* and our everyday notion of truth.

(ii) For me, the main use of KF as a theory of truth is to reason systematically about Kripke's definition of truth (cf. the *Why* reasons). But as I have written in 1984, concerning the Lukasiewicz and Kleene 3-valued logics, "nothing like sustained ordinary reasoning can be carried out in either logic."

II. The System KF (final)

(iii) Still, examples like that of the "Liar" (λ), or the "Revenge of the Liar", may give one pause. Given the consistency of KF, I regard such as marginal "unintended consequences", or "spandrels" in the sense of Stephen Jay Gould. Over all good theories (e.g., Lebesgue measure theory) can have a few bad consequences (e.g., Banach–Tarski Theorem).

III. An Axiomatization of Deflationism using an Intensional Biconditional

- This is a new axiomatization of Deflationism using some old work, "Toward useful type-free theories, I" (1984).
- Notation: S is an extension of PA . S* is a further extension of S using the predicate symbol T(x) together with a new biconditional operator A≡B under which the formulas of L* are closed. The informal interpretation of ≡ is equivalence by definition, even where some instances may not be defined.
- For example, T(A)≡A, or $y \in \{x | A(x)\} \equiv A(y)$.

III. Axiomatization of Deflationism using an Intensional Biconditional (cont'd)

- <u>More notation</u>: Let t be the formula (0=0) and f be its negation. Write D(A) for $(A \equiv t \vee A \equiv f)$. Unless otherwise noted, 'A', 'B',... range over formulas of L(S*). A is called definite or (determinate) if D(A) holds.
- For sentences A, T(A) is T(#A) as before. If A is a formula A(x,y,...) then T(A) is written for T(#A(num.x, num.y,...)).

III. Axioms for Deflationism with ≡ (cont'd)

The system S*:

AX I. $T(A) \equiv A$

AX II. $\equiv$ is an equivalence relation

AX III. $\neg(t \equiv f)$

AX IV. $\equiv$ preserves $\neg$, $\vee$, $\equiv$, and $\forall$

AX V. D(A), for A atomic in L(S).

AX VI. D is strongly compositional w.r.t. $\neg$, $\vee$ and $\forall$.

III. Axiomatization of Deflationism with ≡ (cont'd)

Lemma. S* proves the following:

(i) $D(T(A)) \leftrightarrow D(A)$

(ii) $D(A) \wedge (A \equiv B) \rightarrow D(B) \wedge (A \leftrightarrow B)$

(iii) $D(A) \rightarrow (T(A) \leftrightarrow A)$

(iv) T is strongly compositional w.r.t. definite formulas.

III. Axiomatization of Deflationism with ≡ (cont'd)

<u>Theorem</u> (Feferman 1984) S* is a conservative extension of S.

Two proofs (Feferman and Aczel 1980)

(i) (S.F.) construct a model with a combinatory reduction relation A≥B satisfying Church–Rosser Theorem, then define A≡B to hold if there exists C with A≥C and B≥C.

(ii) (P.A.) Turn the 3-valued Kripke 1975 model into a 2-valued model in an unexpected way.

III. Axiomatization of Deflationism with ≡ (final)

- Since S* is a conservative extension of S, it does not prove L2 (=F2). For example, with S = PA, if S* proved the sentence P expressing that all provable sentences A of PA are true, it would follow that S* proves the consistency of PA.
- Thus, S* is not immune to the "generalization" problem that has been raised for deflationary systems.
- However, by the model construction for S*, the addition of sentences like P to S* preserves consistency if S has standard models.

Some Possible Further Directions of Research

- The Aczel model construction for S* should be amenable to further exploitation, e.g. by adding compositionality axioms for T(x).
- Is there a consistent axiomatization of a *stratified* form of the Tarskian hierarchy, where *stratification* is meant in the sense of Quine's NF (or NFU, consistent by Jensen)? This could be used as a formalization of contextual theories of truth as in Parsons, Burge (in Martin1984).
- What about Gaifman's pointer theory of truth?

References

- P. Aczel and S. Feferman (1980), Consistency of the unrestricted abstraction principle using an intensional equivalence operator, in (J.P. Seldin and J.R. Hindley, eds.) *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, 67-98.
- S. Feferman (1984), Toward useful type-free theories I, in (Martin 1984), 237-287.
- S. Feferman (1991), Reflecting on incompleteness, *J. Symbolic Logic* 56, 1-49.
- S. Feferman (2008), Axioms for determinateness and truth, *Rev. Symbolic Logic* 1, 204-217.
- R. L. Martin, ed. (1984), *Recent Essays on Truth and the Liar Paradox.*