Proof-theoretic remarks on extensions of the Kripke-Feferman theory of truth

Pietro Brocci

SCUOLA NORMALE SUPERIORE

Wormshop 2024
University of Ghent

Part of this presentation is based on joint work with Carlo Nicolai (KCL).

The study of axiomatic theories of truth can be simplified in the following steps:

- Adding a unary predicate for truth to an expressive enough arithmetical base theory (EA, PA).
- Formulating axioms or schemata for the predicate.
- Establishing the amount of new arithmetical theorems provable thanks to truth (i.e., computing the theory's ordinal).

- Axiomatic theories of truth now have many applications in philosophical logic.
- Originally they were introduced as a foundational tool in Feferman project [Feferman, 1991].
- They had the role of defining the reflective closure of an axiomatic system.
- The reflective closure of a system is obtained by transfinitely iterating the addition of reflection principles to the system. The truth predicate makes it possible to dispense of this transfinite process.

- The iteration process reaches a fixed-point at the ordinal Γ_0.
- The corresponding theory of truth is the schematic extension of KF.
- This result fits neatly in Feferman's foundational programme since he understood this ordinal as the limit of predicativity.

More or less recent works challenge that Γ_0 is the limit of predicative mathematics.

[Weaver, 2005] defines a way to predicatively define ordinals up to the Small Veblen Ordinal and suggests that the strategy can be extended to bigger ones.

For these reasons I would like to:

- Find extensions of KF that can reach ”impredicative” strength.
- Compare the foundational role of these theories with the one employed by Feferman.
- Outline independent justifications for the additional principles added to KF (if there are any).

KF

The theory is an axiomatization of Kripkean fixed-point semantics based on Strong Kleene logic. The language $\mathcal{L}_{\mathbb{T}}$ of the theory is obtained by ext. $\mathcal{L}_{\mathsf{PA}}$ with a unary predicate $\mathbb{T}$. Induction is then extended to $\mathcal{L}_{\mathbb{T}}$.

Axiomatically the theory describes a type-free compositional and iterative notion of truth, for example:

$$\forall x \forall y (Sent_{\mathbb{T}}(x \wedge y) \rightarrow (\mathbb{T}(x \wedge y) \leftrightarrow \mathbb{T}x \wedge \mathbb{T}y)$$

$$\forall v \forall x (Sent_{\mathbb{T}}(\forall v x) \rightarrow (\mathbb{T}(\forall v x) \leftrightarrow \forall t \mathbb{T}(x(t/v)))$$

$$\forall t (\mathbb{T}(\mathbb{T}t) \leftrightarrow \mathbb{T}t)$$

Extending the theory

- An extension is already needed by Feferman to obtain a theory with ordinal Γ_0. He adds a principle closely related to the Bar Rule.
- This suggests a generalisation of the role that truth plays:

Truth can play a foundational role because it is a logico-mathematical tool that can simulate second-order talk (s.o. quantification, s.o. variables, ...) in a first-order setting.

Thus, we look for logical notions that yield ”impredicative” strength and employ s.o. talk.

A candidate is a statement that describes the smallest fixed point of an arithmetical operator.

The truth-theoretic version of the principle is Generalised Induction (**GI**) as introduced by [Cantini, 1989]:

$$\forall x(A(x, B) \rightarrow B(x)) \rightarrow \forall x(\mathbb{T}\ulcorner I_A(\dot{x})\urcorner \rightarrow B(x))$$

I want to argue that there are deeper motivations to adopt this principle for **KF**, but I will elaborate at the end.

The Lower Bound

Let $\mathcal{L}_{\mathsf{ID}}$ be the language $\mathcal{L}_{\mathsf{PA}}$ expanded with $\in$ and set constants I_A for any arithmetical formula $A(x, y)$.

The theory ID_1 is obtained from PA by extending induction to $\mathcal{L}_{\mathsf{ID}}$ and the two following principles:

$$\forall x(A(x, I_A) \to x \in I_A)$$

$$\forall x(A(x, B) \to B(x)) \to \forall x(x \in I_A \to B(x))$$

Let $^{\#}$ be a translation from $\mathcal{L}_{\mathsf{ID}}$ to $\mathcal{L}_{\mathbb{T}}$, that preserves the arithmetical statements, commutes with logical operations and interprets set-membership as truth-predication:

$$(t \in I_A)^{\#} = \mathbb{T}\ulcorner I_A(t)\urcorner$$

This makes precise how truth can be a tool to simulate s.o. talk

We can finally verify this. Easily from the definition of the translation one can prove:

Lemma (Cantini)

$\mathsf{KFGI} \vdash (\mathsf{ID}_1)^{\#}$

But it is left to show if the theory's proof-theoretic power exceeds this.

The Upper Bound

The semi-formal system KFGI^{∞} is defined as usual for the arithmetical part with also sequences as syntactic objects. We are drawing from [Pohlers, 1989],[Pohlers, 2008].

Derivations are controlled by Operators (omitted for simplicity).

We rec. define

$$\mathbb{T}^{\prec\zeta}\ulcorner I_A(t)\urcorner := \bigvee_{\eta\prec\zeta} A(\mathbb{T}^{\prec\eta}\ulcorner I_A()\urcorner, t)$$

and

$$\mathbb{T}^{\alpha}\ulcorner I_A(t)\urcorner := A(\mathbb{T}^{\prec\alpha}\ulcorner I_A()\urcorner, t)$$

The system includes

- Rules for $\bigvee$, $\bigwedge$, cut
- **KF** rules for $\mathbb{T}$
- A rule for Ω:

$$\frac{\vdash^{\alpha} \forall x(A(x, \mathbb{T}^{<\Omega}\ulcorner I_A() \urcorner))}{\vdash^{\beta} \mathbb{T}^{<\Omega}\ulcorner I_A(t) \urcorner}\ \mathrm{CL}$$

where Ω can be interpreted as ω_1 or ω_1^{CK}. The rule is needed to have a define Ω, since our semi-formal system cannot contain a recursive ordinal notation system which includes a segment of length Ω.

We define an embedding from **KFGI** that leaves intact truth-predications that do not include I_A.

Proposition

$$\mathsf{KFGI} \vdash \Gamma \Rightarrow \mathsf{KFGI}^{\infty} \left|\frac{\Omega\cdot 2+\omega}{\Omega+n}\right. \Gamma^*$$

From here, it usually goes like this:

$$\mathsf{KF}^{+} \left|\frac{n}{}\right. \varphi \Rightarrow \mathsf{KF}^{\infty} \left|\frac{\omega+n}{m}\right. \Gamma \Rightarrow \mathsf{KF}^{\infty} \left|\frac{\alpha<\varepsilon_0}{0}\right. \Gamma \Rightarrow \models \Gamma[F^{\beta}, T^{\beta+2^{\alpha}}]$$

For $X \subseteq On$ let $\mathcal{H}_\alpha(X)$ be the least set of ordinals containing $\{0, \Omega\}$, which is closed under $\mathcal{H}$ and the *collapsing function* $\Psi_{\mathcal{H}} \restriction \alpha$. The collapsing function is defined as

$$\Psi_{\mathcal{H}}(\alpha) := min\{\xi | \xi \notin \mathcal{H}_\alpha(\emptyset)\}$$

Lemma (Collapsing)

If Γ does not contain Ω-branching conjunctions,
$\mathsf{KFGI}^\infty \left|\frac{\beta}{\Omega+1}\right. \Gamma \Rightarrow \mathsf{KFGI}^\infty \left|\frac{\Psi(\omega^\beta)}{\Psi(\omega^\beta)}\right. \Gamma$

Lemma (Quasi cut-elimination)

$\mathsf{KFGI}^\infty \left|\frac{\Psi(\omega^\beta)}{\Psi(\omega^\beta)}\right. \Gamma \Rightarrow \mathsf{KFGI}^\infty \left|\frac{\Psi(\omega^\beta)}{0}\right. \Gamma$

Asymmetric Interpretation

let A be a formula of $\mathcal{L}_\infty$, we say that $\vDash A[\beta,\alpha]$ holds when:

- Every logical symbol is interpreted in a standard way, except T.
- Let I_Γ^α be the α-th stage of the m.f.p.

$$\vDash Tt[\beta,\alpha] \Leftrightarrow t^{\mathbb{N}} \in I_\Gamma^\alpha$$

$$\vDash \neg Tt[\beta,\alpha] \Leftrightarrow t^{\mathbb{N}} \notin I_\Gamma^\beta$$

$$\vDash \Gamma[\beta,\alpha] := \{A_1[\beta,\alpha], ..., A_n[\beta,\alpha]\}$$

The asymmetric interpretation enjoys a crucial property of persistence:

let $0 < \beta < \beta' < \gamma' < \gamma$, then $\vDash A[\beta',\delta'] \Rightarrow \vDash A[\beta,\delta]$

Lemma

If $\mathsf{KFGI}^\infty \left|\frac{\alpha}{0}\right. \Gamma$ and $\Omega \notin \mathit{par}(\Gamma)$ then for every $\beta > 0$:

$$\vDash \Gamma[\beta, \beta + \Psi(\omega^\alpha))].$$

and by t.i. formalise this in a suitable system $\mathsf{RT}_{<\Psi(\varepsilon_{\Omega+1})}$

Lemma

$\mathsf{RT}_{<\Psi(\varepsilon_{\Omega+1})} \vdash \mathit{Prov}_{\mathsf{KFGI}^\infty}(\ulcorner\Gamma\urcorner, 0, \alpha) \to \mathbb{T}_\beta^{\beta+\Psi(\omega^\alpha)}(\ulcorner\Gamma\urcorner)$

with the lower bound, we can conclude:

$$|\mathsf{KFGI}| = \Psi(\varepsilon_{\Omega+1})$$

But why should we add **GI** and not something else, for instance, Reflection Principles?

If we take our axiom-making process to be a definition of truth, we should keep the same epistemic status as a definition:

> *It is not possible to prove something new from a definition alone that would be unprovable without it. [Frege, 1979]*
>
> *Oughtn't we worry [...] into accepting a substantive metaphysical thesis by insisting that the thesis has the "epistemic status" of a definition? The disquotationalist has a one-word answer: conservativity.[McGee, 2005]*

Proposition (Cantini)

GI is semantically conservative over KF^{-}

Open Problems: categoricity-like features

The theory is proof-theoretically equivalent to KF_μ, see [Burgess, 2014].

[Fischer et al., 2015] argue that $\mathbb{N}$-categoricity is a conceptually relevant feature for axiomatic theories of truth.

KF_μ and likely KFGI do not share this feature.

- As [Enayat and Łełyk, 2024] have recently shown KF_μ enjoys another categoricity-like property that other theories of truth do not, solidity.
- Is this feature shared with KFGI?
- What is the significance of this property for theories of truth?

Thank you!

Bibliography

[Burgess, 2014] Burgess, J. (2014).
Friedman and the axiomatization of kripke's theory of truth.
In Tennant, N., editor, *Foundational Adventures: Essays in Honour of Harvey M. Friedman*. College Publications.

[Cantini, 1989] Cantini, A. (1989).
Notes on formal theories of truth.
Mathematical Logic Quarterly, 35(2):97–130.

[Enayat and Łełyk, 2024] Enayat, A. and Łełyk, M. (2024).
Categoricity-like properties in the first order realm.
Preprint.

[Feferman, 1991] Feferman, S. (1991).
Reflecting on incompleteness.
The Journal of Symbolic Logic, 56(1):1–49.

[Fischer et al., 2015] Fischer, M., Halbach, V., Kriener, J., and Stern, J. (2015).
Axiomatizing semantic theories of truth?

The Review of Symbolic Logic, 8(2):257–278.

[Frege, 1979] Frege, G. (1979).
Logic and Mathematics.
Chicago University Press.

[McGee, 2005] McGee, V. (2005).
Review: Two conceptions of truth? comment.
Philosophical Studies, 124(1):71–104.

[Pohlers, 1989] Pohlers, W. (1989).
Proof theory: An introduction, volume 1407.
Springer Science & Business Media.

[Pohlers, 2008] Pohlers, W. (2008).
Proof theory: The first step into impredicativity.
Springer Science & Business Media.

[Weaver, 2005] Weaver, N. (2005).
Predicativity beyond gamma_0.

arXiv preprint math/0509244.