Numbers and Truth

The philosophy and mathematics of arithmetic and truth

A Marcus Wallenberg Symposium

19-21 October 2012, Gothenburg

Book of abstracts

Welcome to the conference 'Numbers and Truth', organized by the Department of Philosophy, Linguistics and Theory of Science at the University of Gothenburg in cooperation with the Institute of Philosophy at the University of Warsaw and the Department of Philosophy at the Lund University.

We would like to thank our sponsors:

- Riksbankens Jubileumsfond,
- Marcus Wallenbergs Stiftelse för Internationellt Vetenskapligt Samarbete,
- Wenner-Gren Foundations,
- the Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg.

Warm thanks go also to our scientific committee: Staffan Angere, Sebastian Enqvist, Martin Fischer, Volker Halbach, Leon Horsten, Martin Kaså, Richard Kaye, Juliette Kennedy, Leszek Kołodziejczyk, Roman Kossak, Carlo Proietti, Rafał Urbaniak, Albert Visser, Sean Walsh, and Konrad Zdanowski.

Organizers

Cezary Cieśliński (University of Warsaw)
Fredrik Engström (University of Gothenburg)
Paula Quinon (Lund University)

Contents

Abstracts for contributed talks

Special session

Preface

Natural numbers belong to the most commonly known mathematical entities and are studied and discussed from various angles. Philosophers, mathematicians, cognitive scientists and computer scientists conduct research having for its aim a deeper understanding of these objects. The present conference 'Numbers and Truth' will contribute to these efforts.

Within recent philosophy of mathematics a couple of competitive philosophical schools came to the foreground; importantly, structuralism and neo-Fregeanism. Both of them espouse realism about mathematical objects: these objects exist independently of our minds. In addition, both espouse realism about truth: arithmetical sentences have objective truth values. In this context a natural question arises whether theories of arithmetical truth provide us a better understanding of what natural numbers are. What is the relation between a choice of a particular truth theory for arithmetical language and one's overall standpoint in the philosophy of mathematics, or maybe even one's preferred way of doing mathematics? On a general level, the aim of the conference is to address issues of this sort.

Invited speakers: Martin Fischer, Volker Halbach, Leon Horsten, Richard Kaye, Juliette Kennedy, Roman Kossak, Rafał Urbaniak, Albert Visser, Sean Walsh, Konrad Zdanowski.

Contributed speakers: Walter Dean, Hidenori Kurokawa, Graham E. Leigh, Edoardo Rivello, Andrea Strollo, Shunsuke Yatabe.

Timetable

Friday, 19 October

- 13:30-14:30 Welcome/Registration/Coffee
- 14:30-15:30 Leon Horsten: *Truth, conditionals, and paradox*
- 15:45-16:45 Juliette Kennedy: *Change the logic, change the meaning?*
- 16:45-17:00 Coffee
- 17:00-17:45 Walter Dean and Hidenori Kurokawa: *Numbers, constructive truth, and the Kreisel-Goodman paradox*
- 17:50-18:35 Andrea Strollo: *The disentanglement of syntax from a model theoretic point of view*
- 18:40-19:25 Graham E. Leigh: *A proof-theoretic account of classical principles of truth*
- 20:00 Dinner at Restaurant Tara's

Saturday, 20 October

- 9:30-10:30 Roman Kossak: *Model theory of satisfaction classes*
- 10:30-11:00 Coffee
- 11:00-12:00 Albert Visser: *Degrees of interpretability of finitely axiomatized sequential theories*
- 12:15-13:15 Richard Kaye: *Adding standardness to nonstandard models*
- 13:15-14:30 Lunch
- 14:30-15:30 Rafał Urbaniak: *Neologicism: for real(s)?*
- 15:45-16:45 Konrad Zdanowski: *On notations systems for natural numbers and polynomial time computations*
- 16:45-17:15 Coffee
- 17:15-18:00 Shunsuke Yatabe: *Yablo's paradox and ω-inconsistency*
- 18:15-19:00 Edoardo Rivello: *Revision without ordinals*
- 19:30 Dinner at Restaurant Familjen

Sunday, 21 October

- 9:30-10:30 Sean Walsh: *Empiricism, probability, and knowledge of arithmetic*
- 10:30-11:00 Coffee
- 11:00-12:00 Martin Fischer: *Truth and Groundedness*
- 12:15-13:15 Volker Halbach: *Axiomatic and semantic approaches to truth*
- 13:15-14:30 Lunch
- 14:30-16:30 Special session, Roman Kossak: *Husserl's philosophy of arithmetic, a discussion*

Practical information

Venues

The conference will take place at the Department of Philosophy, Linguistics and Theory of Science, at Olof Wijksgatan 6, Göteborg.

The recommended hotel is Hotel Panorama, it is located at Eklandagatan 51-53. It is a short 10 minutes walk from the conference venue.

Lunches and dinners

Lunches and dinners will be free for speakers. Lunch will be served at the conference venue. Dinner on Friday will be at Resturant Tara's at Södra vägen 24, and on Saturday at Restaurant Familjen at Arkivgatan 7.

Wireless internet access

Wireless internet access will be available to all participants though a guest account on the Wireless network with SSID GöteborgsUniversitet. The login is web-based. Eduroam will also be available to everyone with the appropriate account.

Truth and Groundedness

Martin Fischer
MCMP, Munich

At least since Kripke the notion of groundedness has played a major role in attempts to solving the semantic paradoxes. In Kripke's account the notion of groundedness is explicated via his fixed-point construction and especially the minimal fixed-point. In Yablo's terminology Kripke's work is focused on the inheritance aspect of groundedness and not on the dependence aspect.

In the talk I will present a variation of Kripke's fixed-point construction that is a model for a theory of truth and groundedness. The aim of this model is twofold: On the one hand it tries to capture the structure of the fixed-points. On the other hand it provides a justification for a theory based on a quite expressive language, a language containing not only a truth predicate but also a groundedness predicate.

The model is a fixed-point construction combined with a possible worlds model. The construction in the talk is based on work carried out by Volker Halbach and Philip Welch as well as Johannes Stern. The innovation of my construction is that in the end each possible world represents one possible fixed-point which allows for an adequate representation of grounding.

The language $\mathcal{L}_G$ is the language of arithmetic expanded by two one place predicates T, G. A model $\mathfrak{M} = (W, R, f)$ with W a set of worlds R an accessibility relation and f an evaluation function that assigns to each world an extension of the truth predicate, i.e. a consistent set of sentences. The extension of the groundedness predicate is dependent on the extension of the truth predicate. The extension of G at a world w will be the set of those sentences for which either they are true at all accessible worlds or their negation is. With the help of the strong Kleene scheme it is possible to define a monotone jump operator that guarantees the existence of a fixed-point. By starting with a suitable model the fixed-point has some nice properties. Moreover in the closed-off versions some interesting theorems hold and point to possible theories of truth and groundedness.

The talk will contain a discussion of the adequacy of the model to capture our intuitions about groundedness and possible shortcomings. One of the advantages of the model is its flexibility to accommodate different intuitions of grounding, for example following different evaluation schemes. Moreover we can use the model to introduce further notions such as paradoxicality or intrinsically true.

Axiomatic and semantic approaches to truth

Volker Halbach
University of Oxford

Axiomatic and semantic approaches can be found already in Tarski's seminal work on formal theories of truth. Syntactic or axiomatic principles have been used in order to evaluate, motivate, and justify semantic definitions of truth; conversely, semantic constructions and analyses have been used for similar purposes with respect to axiomatic approaches.

In the talk I will try to analyze the interplay between axiomatic and semantic theories of truth. In particular, I will argue that Tarski's rejection of axiomatic theories of truth is incoherent with his own theory of truth.

Since Donald Davidson's work on truth and meaning, it has often been claimed that certain semantic theories of truth can be captured by or rephrased in an axiomatic theory. I will try to analyze what this could mean, not only for classical typed theories such as Tarski's but also for type-free theories such as Kripke's.

Truth, conditionals, and paradox

Leon Horsten
University of Bristol

Field takes Kripke's theory of truth to be deficient in two ways:

1. Kripke's theory does not make the unrestricted Tarski-biconditionals come out true
2. Kripke's theory does not contain a real conditional

In order to address these problems, Field interleaves the Kripkean fixed point construction with a revision construction in an ingenious way.

Yablo has proposed an alternative theory of truth and conditionals that has a uniform (Kripkean) conceptual motivation. However, as Field pointed out, Yablo's theory does not give a sensible account of embedded conditionals.

In my talk, I want to elaborate Yablo's suggestion in a way that enables embedded conditionals to be treated in a more satisfactory manner.

ADDING STANDARDNESS TO NONSTANDARD MODELS

Richard W. Kaye
University of Birmingham

If M is a nonstandard model of PA, the set of standard natural numbers $\omega = \{0, 1, 2, \ldots\}$ forms an initial segment of M in a canonical way. It therefore makes sense to look at the model (M, ω) formed by adding a predicate for ω to M. Model theory of (M, ω) is rather hard. Such structures encode a number of very difficult questions in second order number theory (analysis), including many open ones. However some attempt can be made to understand (M, ω) relative to second order schemes of arithmetic.

The talk will touch on three related questions: what can one interpret or define in (M, ω) that cannot be defined in M; what is the theory of structures such as (M, ω); and what reals are coded in the model (M, ω)?

In terms of definability and interpretability, structures (M, ω) interpret ω-models of second order arithmetic, and also under certain circumstances define truth predicates for submodels of M. Thus second order systems motivate this work on structures (M, ω). It turns out that the truth predicate has useful applications and some details of how it arises will be given.

In terms of the theory, the Henkin-Orey theorem on ω-logic tells us about the theory of all models (M, ω) (i.e. the statements true in all such models) but tells us little about the theory of any specific model. In fact the theory of (M, ω) depends on structural properties of M that are not first order, and so there is a wide range of possibilities for the theory of (M, ω), even for a given completion T of PA. It is perhaps surprising therefore that given a complete theory T extending PA there is a canonical choice for a theory $\mathrm{Th}(M, \omega)$ of some $M \models T$. More surprisingly, this is not hard to prove. We will discuss a few consequences of this result, and further applications of the truth predicates available in (M, ω) will be given.

In a difficult paper in the JSL, Kanovei characterised the Scott sets $\mathrm{Rep}(M, \omega)$ of subsets of ω that are 0-definable in (M, ω), when M is a model of true arithmetic. A similar characterisation of the standard system $\mathrm{SSy}(M, \omega)$ of (M, ω) (i.e. such definable sets, where parameters are allowed from M) is not known. We will conclude with some results and observations on these standard systems, with some open problems for future work.

Much of this is joint work with Roman Kossak and Tin Lok Wong.

Change the logic, change the meaning?

Juliette Kennedy
University of Helsinki

In his 1946 Princeton Bicentennial Lecture Gödel suggested the problem of finding a notion of definability for set theory which is ”formalism free” in a sense similar to the notion of computable function — a notion which is very robust with respect to its various associated formalisms. One way to interpret this suggestion is to consider standard notions of definability in set theory, which are usually built over first order logic, and change the underlying logic. We show that constructibility is not very sensitive to the underlying logic, and the same goes for hereditary ordinal definability (or HOD). We observe that under an extensional notion of meaning for set theoretic discourse, Quine's Dictum ”change of logic implies change of meaning” is only partially true. This is joint work with Menachem Magidor and Jouko Väänänen.

Model theory of satisfaction classes

Roman Kossak
City University of New York

All countable recursively saturated models of Peano Arithmetic have nonstandard satisfaction classes. In fact, each such model has a great variety of nonstandard satisfaction classes. I will survey model theoretic techniques that can be applied to construct many different inductive satisfaction classes, and I will show how, in return, inductive satisfaction classes are used to prove important result about recursively saturated models of PA. I will also pose an open problem concerning a possible converse to Tarski's undefinability of truth theorem.

Neologicism: for real(s)?

Rafał Urbaniak
University of Gdańsk
Ghent University
Trinity College Dublin

Neologicism to be successful should provide foundations not only for natural number arithmetic, but also for other mathematical theories, like real number theory (RNT). The search for appropriate abstraction principles is tricky. They should be strong enough to give the desired theory, not strong enough to prove undesired claims and (last but not least) provide basis for a philosophically acceptable story about RNT. I will survey existing attempts of developing neologicist foundations of RNT, evaluate them and try to improve on them.

Degrees of interpretability of finitely axiomatized sequential theories

Albert Visser
Utrecht University

Finitely axiomatized sequential theories are something like a natural kind of theories. They share a lot of salient and important properties. Moreover, many familiar theories belong to this kind. Examples of finitely axiomatized sequential theories are the basic theory PA^-, Buss's theory S^1_2, Elementary Arithmetic EA, $\mathrm{I}\Sigma_1$, ACA_0 and the Gödel-Bernays theory of sets and classes GB.

The study of interpretability degrees of a class of theories is important as the study of a notion of the strength of a theory. For example the Observed Linearity of Reverse Mathematics only comes into focus against the background of the result that the surrounding degree structure contains infinite anti-chains.

In this talk we give an introduction to the interpretability degrees of finitely axiomatized sequential theories. We are especially interested in the question: how are the degrees of extensions (in the same language) of a given theory embedded in the complete degree structure? We will briefly look at the case of a non-sequential theory, to wit: Robinson's Arithmetic Q. This case shows interesting similarities and differences with the sequential one.

As we will see arithmetical theories play a central role in the study of the interpretability degrees of finitely axiomatized sequential theories. This is already visible in the classical result that each such degree contains an arithmetical theory.

Empiricism, probability, and knowledge of arithmetic

Sean Walsh
University of California, Irvine

In this talk, the tenability of extending arithmetical knowledge by way of confirmation is examined, where the relevant notion of confirmation is understood probabilistically in the manner familiar from Bayesianism. The motivation here is to see what can be said for a pre-Fregean view to the effect that mathematical induction– one of the Peano axioms– is akin to enumerative induction in certain of its epistemic features. I will focus on one ostensible problem with this view, namely that from certain perspectives the arithmetical probabilities in question seem just as intractable as arithmetical truth itself, either because of the inherent complexity of the probabilities in question as measured by the methods of computability theory, or because the most obvious examples of arithmetical probabilities in this sense are counting measures and hence in essence are just weighted averages of arithmetical truths.

On notations systems for natural numbers and polynomial time computations

Konrad Zdanowski
Cardinal Stefan Wyszyński University, Warsaw

We have many notational systems for denoting natural numbers: unary, binary, decimal, p-adic, residues modulo a given sequence of numbers and many others. Indeed, different notations may be well suited for different purposes. Also the structure of a computing device that one uses may favour one notation system over the other one (e.g. computing in a modulo residues notation is well suited for parallel computations).

Thus, there is not a single "best" notation. However, we can investigate the set of all possible notations for natural numbers and investigate their general properties. In the talk I will show that the decimal notation (or, rather, its equivalence class) is a maximal notation in a certain natural ordering between feasible notations.

During the second part, I will talk about polynomial time recursion schemes for the notation for hereditarily finite sets in a style of Bellantoni and Cook (which is a work in progress).

Numbers, constructive truth, and the Kreisel-Goodman paradox

Walter Dean
University of Warwick

Hidenori Kurokawa
City University of New York

The historical development of intuitionism is often presented, at least in part, as a reaction to the set theoretic paradoxes. For instance, it was in light of Richard's Paradox that Poincaré originally proposed that predicative restrictions be imposed on set existence, a reaction which Bouwer [3] would also later articulate. But considerably less attention has been paid to the status of the semantic paradoxes, such as the Liar or Curry's Paradox, in regard to intuitionism. This is in some sense unexpected, however, since many expositors (e.g. Dummett [6], Prawitz [13]) have suggested that the notion of truth or validity relevant to constructive mathematics should be understood to satisfy a principle of the form

(1) $A \leftrightarrow \exists p(p \text{ is a constructive proof of } A)$

As an attempt to characterize a conception of truth, however, (1) immediately stands out in virtue of its similarity to the T-schema (i.e. $A \leftrightarrow \mathrm{Tr}(\ulcorner A \urcorner)$), the adjunction of which to formal arithmetic leads to Tarski's formalized version of the Liar Paradox. Although this result is commonly cited as a fundamental antinomy of the naive conception of truth, it need not be taken as a truly *mathematical* antinomy. For instance, we may conclude along with Tarski that it merely shows that no sufficiently expressive mathematical language can be semantically closed (i.e. capable of expressing its own truth predicate).

On the other hand, a variety of proof theoretic techniques (e.g. Kleene realizability, the Curry-Howard correspondence) can be understood as attempting to analyze the notion of constructive proof so as to make good on the intention expressed by (1). Inasmuch as these are developments *within* constructive mathematics, one might reasonably fear that the adoption of (1) will lead to a paradox not just of the "naive" notion of constructive truth (which, at least *prima facie*, one might think was also unanalyzable within a sufficiently expressive intuitionstic object language), but of constructive mathematics itself. (For since familiar meta-mathematical results like the Diagonal Lemma are derivable intuitionistically, one might think that a formal analysis of proof which validated (1) would lapse into inconsistency in the same manner as the Liar.)

Although such a concern has occasionally been voiced explicitly (e.g. [14]), a more common reaction among intuitionists is to deny that a *formalized* version of (1) should be accepted relative to an appropriate understanding of the intuitionistic connectives. For instance, Beeson [2] remarks that to accept the left-to-right direction of (1) is to acknowledge that there is a fixed universe of proofs to which the constructive existential quantifier may be meaningfully applied. And to do so seems counter to the traditional observation that the notion of constructive provability ought to be regarded as open-ended or "indefinitely extensible".

Note, however, that if we write $P(\ulcorner A \urcorner)$ to express "there exists a constructive proof of A", the following principle still seem unobjectionable:

(T) $P(\ulcorner A \urcorner) \to A$

(Nec) If $\vdash_{\mathsf{Z}} A$, then $\vdash_{\mathsf{Z}} P(\ulcorner A \urcorner)$

Relative to the intuitionistic interpretation of the existential quantifier, T records the fact that if we possess a constructive proof p of A, we can conclude that A is true simply by examining p itself. On the other hand, when Z is taken to subsume some intuitionistic formal system such as Heyting Arithmetic (HA) together with T, then Nec reflects the fact that formal derivability within such a system, while not necessarily exhausting constructive provability, is at least faithful to basic constructive principles.

As benign as these principles may seem, the system consisting of Z, T, and Nec can be shown to be inconsistent by the following intuitionistically valid derivation:

(2)		
	i) $\vdash_{\mathsf{Z}} D \leftrightarrow \neg P(\ulcorner D \urcorner)$	Diagonal Lemma for Z
	ii) $\vdash_{\mathsf{Z}} P(\ulcorner D \urcorner) \to D$	T
	iii) $\vdash_{\mathsf{Z}} \neg P(\ulcorner D \urcorner)$	i), ii)
	iv) $\vdash_{\mathsf{Z}} D$	i), iii)
	v) $\vdash_{\mathsf{Z}} P(\ulcorner D \urcorner)$	Nec, iv)
	vi) $\vdash_{\mathsf{Z}} \bot$	iii), v)

This derivation was first presented by Myhill [12] in the context of discussing the concept of "absolute" provability in light of Gödel's proposal that the Heyting (or "BHK") interpretation of the intuitionistic connectives can be formalized by an embedding into the modal logic S4 – an observation which serves as the basis of Leitgeb's [11] dubbing of (2) as the "Provability Liar".

Our first goal in this note will be to illustrate how consistency concerns closely related to the Provability Liar arose independently within another significant (but less well known) investigation of the Heyting interpretation – namely, the so-called *Theory of Constructions* ($\mathcal{C}$), originally formulated by Kreisel [10], and further refined by Goodman [8], [9]. More generally, however, our aim will be to suggest that this system and its intended interpretation are of contemporary interest in light of the following observations:

I) The Provability Liar is prototypical of a class of inconsistency results wherein the left-to-right direction of the T-schema is dropped in favor of a family of principles (of which Nec is paradigmatic) which allow for the internalization of reasoning carried out within a formal theory of truth or provability.[1] When reconstructed in a theory such as $\mathcal{C}$, it is possible to discern additional logical structure in the use of Nec, thus opening new avenues for responding to paradoxes in this family.

II) Traditional treatments of self-referential truth and provability typically proceed through an explicit arithmetization of syntax, performed over a background theory such as PA or HA. However, Goodman's goal was to show that the meaning of the intuitionistic connectives themselves could be analyzed in a yet more primitive formalism which is intended to be "type- and logic-free". Thus rather than starting from an arithmetical base system, $\mathcal{C}$ is developed by adding a proof operator

[1] This includes several of similar inconsistency results reported by Friedman and Sheard [7], as well as the so-called ω-inconsistency theorem of McGee.

π to pure combinatory logic (or equivalently – and as we shall present it here – untyped lambda calculus). As was first observed in the formulation of the Curry Paradox [4], self-reference arises in this setting without arithmetization due to the existence of so-called *fixed point combinators*. Since an appropriate analogue of Nec is a derivable rule of Goodman's system, a contradiction follows by assuming that the proof-theorem relation is decidable, in conjunction with a version of the reflection principle T.

III) Goodman also proposed a means of interpreting HA within a variant of $\mathcal{C}$ in which constructions are stratified into so-called *grasped domains*. This approach is motivated by concerns arising from the Provability Liar about the surveyablity of the class of *all* constructive proofs. This development is also of interest because it inverts the traditional order of analysis employed in formal treatments of truth and provability – i.e. rather than using a number theoretic language to formulate a truth or provability predicate by arithmetization, Goodman suggests that it is possible to define the natural numbers in terms of putatively more fundamental principles about the combinatory structure of constructive proofs.

We will now proceed by sketching a streamlined version of the Provability Liar in $\mathcal{C}$, and the rudiments of Goodman's proposed interpretation of arithmetic. We will then briefly record some additional questions which we will explore more fully in the sequel.

The system $\mathcal{C}$ and the Kreisel-Goodman Paradox

$\mathcal{C}$ can be most simply presented as an equational calculus (with identity $\equiv$) consisting of variables $x, y, z, \ldots$ intended to denote constructive proofs as well as symbols for the operators required to formalize the Heyting interpretation. Among these are a function abstraction operator (λ), and pairing and projection operators D, D_1, D_2 such that Dxy denotes the pair xy and $D_i(x_1x_2) = x_i$. Additionally, we introduce constant symbols $\top, \bot$ to denote truth and falsity, the traditional combinator $\mathsf{K}xy = x$, and also a primitive binary *proof operator* π with intended interpretation

(3) $\pi uv \equiv \top$ just in case v is a proof that $uz \equiv \top$ for all z

On the intended interpretation of $\mathcal{C}$, $\pi uv \equiv \top$ is understood to be *decidable*. This reflects the traditional intuition that we ought to be able to determine by inspection whether v is a proof that u is satisfied by z.[2]

The foregoing stipulation has a significant consequence for the formalization of the proof condition assigned to intuitionistic implication – i.e. *a proof of $\varphi \to \psi$ consists of a construction x such that if z is a proof φ then xz is proof of ψ, together with a proof which verifies that this construction always works.* As the latter statement is itself a conditional, this analysis might at first be understood as providing a circular characterization of intuitionistic implication. However, on the assumption that v *is a proof of* φ corresponds to a decidable statement, it is possible to render this conditional as

$$(\varphi \to \psi)^* = \lambda xy.\pi(\pi(\varphi^*)x \supset_1 \lambda z.\pi(\psi^*)x(D_2yz))(D_1y)$$

where $\supset_1$ denotes *classical* implication[3] and $(\cdot)^*$ denotes Goodman's proposed embedding of intuitionistic logic into $\mathcal{C}$.

[2] In both form and motivation, $\pi uv \equiv \top$ can thus be compared to statements of the form $\text{Proof}(\overline{n}, \ulcorner\varphi(x)\urcorner)$ where $\text{Proof}(x, y)$ is a standard arithmetical proof predicate.

[3] If we take $\top =_{df} \lambda xy.x$ and $\bot =_{df} \lambda xy.y$, then such an connective is definable is $\mathcal{C}$ as $\lambda xyz.xzy(\lambda w.\top)z$.

Now note that within $\mathcal{C}$ it is possible to formalize

(4) w does not prove that $yz \equiv \top$

as $(\pi y \supset_1 (\lambda x.\bot))w \equiv \top$. If we now let $h =_{df} \pi(\pi(y \supset_1 \lambda x.\bot))(zz)$ then $hyz \equiv \top$ asserts that zz is defined and is a proof that (4) is unprovable. A self-referential statement about provability may now be obtained by considering the fixed-point combinator $a =_{df} \lambda z.(\lambda y.h(yy)z)(\lambda y.h(yy)z)$ and noting that $\vdash az \equiv haz$ – i.e. the statement

(5) $az \equiv \top$

holds just in case zz is a proof that no x is a proof of (5). If we understand $\bot \equiv \top$ to express a contradiction in $\mathcal{C}$, then the following rules which are respectively intended to formalize the decidability (i.e. truth evaluabilty) of πuv and what we will refer to as an *explicit reflection principle* – i.e. that the existence of a specific proof of a statement, entails its truth:

(Dec) If $\pi uv \equiv \top \vdash \bot \equiv \top$, then $\vdash \pi uv \equiv \bot$ (ExpRef) $\pi uv \equiv \top \vdash uz \equiv \top$

It may also be shown the $\mathcal{C}$ has the ability to internalized its own proofs in the sense that

(Int) If $\vdash u \equiv \top$, then there exists some construction v such that $\vdash \pi uv \equiv \top$.

is a derivable principle in which the term v may be constructed from the proof of $u \equiv \top$ in $\mathcal{C}$.

The Kreisel-Goodman Paradox corresponds to the following derivation in $\mathcal{C}$ (we present a parallel derivation in the arithmetical system Z as a gloss):

i)		$\vdash az \equiv haz$	defn. of a	$\vdash A \leftrightarrow P(\ulcorner \neg P(\ulcorner A \urcorner) \urcorner)$	Diag. Lemma
ii)	$\pi(az)x \equiv \top$	$\vdash az \equiv \top$	ExpRef	$P(\ulcorner A \urcorner) \vdash A$	T
iii)	$\pi(az)x \equiv \top$	$\vdash haz \equiv \top$		$P(\ulcorner A \urcorner) \vdash P(\ulcorner \neg P(\ulcorner A \urcorner) \urcorner)$	
iv)	$\pi(az)x \equiv \top$	$\vdash \pi(\pi(az) \supset_1 (\mathsf{K}\bot))(zz) \equiv \top$	defn. of h		
v)	$\pi(az)x \equiv \top$	$\vdash (\pi(az) \supset_1 (\mathsf{K}\bot))w \equiv \top$	ExpRef		
vi)	$\pi(az)x \equiv \top$	$\vdash (\mathsf{K}\bot)w \equiv \top$		$P(\ulcorner A \urcorner) \vdash \neg P(\ulcorner A \urcorner)$	
vii)	$\pi(az)x \equiv \top$	$\vdash \top \equiv \bot$			
viii)		$\vdash \pi(az)x \equiv \bot$	Dec	$\vdash \neg P(\ulcorner A \urcorner)$	
ix)		$\vdash (\pi(az) \supset_1 (\mathsf{K}\bot))x \equiv \top$	defn. $\supset_1$		
x)		$\vdash \pi(\pi(az) \supset_1 (\mathsf{K}\bot))fx \equiv \top$	Int	$\vdash P(\ulcorner \neg P(\ulcorner A \urcorner) \urcorner)$	Nec
xi)		$\vdash \pi(\pi(az) \supset_1 (\mathsf{K}\bot))ff \equiv \top$	subst. for f for x		
xii)		$\vdash haf \equiv \top$	defn. of h		
xiii)		$\vdash af \equiv \top$	defn. of a	$\vdash A$	
xiv)		$\vdash \pi(af)b \equiv \top$	Int	$\vdash P(\ulcorner A \urcorner)$	Nec

As x and z are free at line vii), a contradiction now follows from lines viii) and xiv).

Goodman proposed to resolve the paradox by replacing $\mathcal{C}$ with a theory $\mathcal{C}^*$ in which axioms are adopted which attempt to formalize the stratification of constructions into a cumulative hierarchy $L_0, L_1, \ldots$ such that L_{i+1} is formed by reflecting on the operation of constructions available at L_i and seeing that they are defined on all of their arguments. But he also asserts that it is not in keeping with constructive principles to assume that we "understand the notion of proof in an absolute sense . . . [so as to] visualize the entire constructive universe" [9], p. 14. On this basis he introduces a ternary proof operator $\pi xyz \equiv \top$ which receives the intended interpretation x *is a grasped domain containing* y *and* z *is a proof that* $yw \equiv \top$ *for all* w *in* x. If we now attempt to reformulate the paradox using such an operator, we must assume that there is a level n in which the proof corresponding to the outer occurrence of π at step iv) inhabits. However, Goodman asserts that the inference from x) to xi) will now be unjustified as the term f constructed by internalizing the prior steps will be of level $n + 1$.

Interpreting Heyting Arithmetic

The availability of a pairing operator on proofs is explicit in the Heyting interpretation – e.g. a proof of a conjunction $\varphi \wedge \psi$ is stipulated to be a pair of constructions xy such that x is a proof of φ and y is a proof of ψ. Goodman proposes that a pairing operator can be used to provide a definition of the natural numbers as follows: 0 is interpreted as the combinator K and if n is interpreted as a construction u, then $n+1$ is interpreted as the pair $D(\mathsf{K}u)$. On this basis, he shows how to formulate a decidable natural number predicate in $\mathcal{C}^*$ for which an appropriate induction rule holds. The interpretation of HA itself is accomplished by using $\mathcal{C}^*$ to analyze the logical connectives in accordance with the Heyting interpretation. (We postpone the details.)

Some questions

1) Goodman's proposed means of responding to the Provability Liar is in some sense reminiscent of typed theories of truth in the style of Tarski's hierarchy of metalanguages. In particular, Goodman stresses that his proposed stratification of constructions is one not of *logical type*, but rather one about the *subject matter* of mathematical proofs. Is an approach which splits the "naive" notion of constructive provability into levels more plausible than proposals which make a similar distinction about the "level" at a which a sentence may be asserted to be true in the classical setting? How does this approach compare to Anderson's [1] resolution of the structurally similar Paradox of the Knower based on a hierarchy of knowledge predicates?

2) By treating proofs as freestanding objects, theories like $\mathcal{C}$ refine approaches which represent truth or provability as a predicate. However, $\mathcal{C}$ does not explicitly represent the quantifier over proofs expressed by (1) or (implicitly) in the fixed point statement which appears in the Provability Liar. What first-order principles about proofs are subsumed by $\mathcal{C}$ and $\mathcal{C}^*$? How do these principles figure in the construction of the terms f and b which figure in the derivation of the Kreisel-Goodman Paradox? Can such principles be used to make precise the intuition that the notion of constructive proof is indefinitely extensible?

3) Goodman's embedding of HA into $\mathcal{C}^*$ is one of several methods which have been proposed for interpreting first-order logic and arithmetic in combinatory logic. Although the Curry Paradox was originally presented using such a theory, systems of this sort also share an important affinity with *formalism* in the philosophy of mathematics (e.g. in the sense developed by Curry in [5]) in that they seek to analyze numerical and logical notions in terms of purely combinatory ones. As systems of combinatory logic are often presented as *uninterpreted*, it seems misleading to describe them as taking for a granted a substantial notion of truth (e.g. truth in the standard model). What, then, should be said about the conceptual significance of formal inconsistency results like the Provability Liar in this setting? And what does this tell us more generally about the relationship between arithmetic, truth, and self-reference?

Bibliography

[1] C. A. Anderson. The paradox of the knower. *The Journal of Philosophy*, 80(6):338–355, 1984.

[2] M. Beeson. *Foundations of constructive mathematics: metamathematical studies*. Springer-Verlag, 1985.

[3] L. Brouwer. Intuitionism and formalism. In P. Benaceraff and H. Putnam, editors, *Philosophy of Mathematics: Selected Readings*, pages 77–89. Cambridge University Press, 1983.

[4] H. Curry. The inconsistency of certain formal logic. *Journal of Symbolic Logic*, pages 115–117, 1942.

[5] H. Curry. *Outlines of a Formalist Philosophy of Mathematics.* North-Holland, Amsterdam, 1951.

[6] M. Dummett. Realism. *Synthese*, 52(1):55–112, 1982.

[7] H. Friedman and M. Sheard. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33:1–21, 1987.

[8] N. Goodman. A theory of constructions equivalent to arithmetic. In J. M. A. Kino and R. Vesley, editors, *Intuitionism and Proof Theory*, pages 101 – 120. Elsevier, 1970.

[9] N. Goodman. The arithmetic theory of constructions. *Cambridge Summer School in Mathematical Logic*, pages 274–298, 1973.

[10] G. Kreisel. Foundations of intuitionistic logic. *Studies in Logic and the Foundations of Mathematics*, 44:198–210, 1962.

[11] H. Leitgeb. On formal and informal provability. In *New Waves in the Philosophy of Mathematics.* Palgrave Macmillan, 2009.

[12] J. Myhill. Some Remarks on the Notion of Proof. *The Journal of Philosophy*, 57:461–471, 1960.

[13] D. Prawitz. Truth and Objectivity from a Verificationist Point of View. In G. Dales and G. Oliveri, editors, *Truth in Mathematics*, pages 41–51. Oxford University Press, 1998.

[14] N. Weaver. Intuitionism and the Liar Paradox. *Annals of Pure and Applied Logic*, 2012.

A proof-theoretic account of classical principles of truth

Graham E. Leigh
University of Oxford

In the analysis of truth it is often logic that is blamed for the truth-theoretic paradoxes. It is well known that compositionality and self-referentiality do not mix well in a purely classical environment, so logics weaker than classical logic have been suggested for the truth predicate. These typically involve moving to partial or paraconsistent logics (see for example [3, 6]). Compared with classical logic, such logics are not well understood, and as Feferman observes, it is questionable whether "[anything] like sustained ordinary reasoning can be carried [out in them]" [2, p. 95]. In this paper we seek to explore more deeply the interface between principles of self-applicable truth and classical logic, in particular the rôle classical principles play in restricting the free use of truth principles. We propose to do this by analysing various axiomatisations of truth over intuitionistic logic.

Intuitionistic logic has not received much attention to date from the truth theory community, but it does have a number of virtues. As a logic upon which to study the effect of classical reasoning, it is better suited than other weakenings of classical logic because of its mature model theory and proof theory. This means that consistency and conservativeness results for truth can be easily interpreted outside the field of theories of truth (for example in set theory or second order analysis) as they can with a classical base theory. It also provides a firm base on which to study constructive interpretations of truth. Of course few authors, if any, believe that truth is inherently constructive. But since an axiomatic theory of truth can provide at best only an approximation for "real" truth, it is natural at least not to rule out constructive formulations.

The analysis we present is inspired by the work of Friedman and Sheard [4]. There the authors determine the consistency of all subsets of a collection of twelve principles of truth over a classical base theory. The outcome is nine maximal consistent theories of truth. Subsequent work by Cantini [1], Halbach [5] and Leigh and Rathjen [8] establish the proof-theoretic strength of each of these theories showing they range from conservative extensions of Peano arithmetic to the strength of one inductive definition, ID_1. These results are summarised in table 2 below. Beginning with [7] the work has been transferred to intuitionistic logic. It is shown that some of the inconsistencies noted in [4] can be attributed to classical principles inherent in the base theory, the law of excluded middle or the statement that the logic under the truth predicate is classical. Moreover, these sets of principles turn out to be consistent over a purely intuitionistic base theory. For example, over the classical base theory utilised in [4], the principles $\mathtt{T}(A \vee B) \rightarrow \mathtt{T}A \vee \mathtt{T}B$ and $\mathtt{T}A \vee \mathtt{T}(\neg A)$ (stated for arbitrary sentences A, B) are equivalent; nevertheless, there are models of the intuitionistic base theory that satisfy the former but refute the latter principle. In fact, the former axiom is consistent with all consistent sets of truth principles and the latter is inconsistent with roughly half. The upshot is that although there are still exactly nine maximal consistent sets of principles over the new base theory, reverting to intuitionistic logic provides more freedom to express principles of truth while avoiding the pitfalls of inconsistency.

In the present paper we discuss the proof-theoretic analysis of the induced maximal consistent theories. Four of the theories remain sub-theories of their classical counterparts. The other theories are inconsistent with either the law of excluded middle or the principle of a classical truth predicate (or both). Upper bounds on their proof-theoretic strength do not, therefore, immediately result from the analysis of their classical cousins and a more refined analysis is required.

We denote by $\mathcal{L}_{\mathtt{T}}$ the language of arithmetic, $\mathcal{L}$, extended to include a unary predicate $\mathtt{T}$. $\mathsf{HA}_{\mathtt{T}}$ is Heyting arithmetic formulated in $\mathcal{L}_{\mathtt{T}}$, that is with the schema of induction extended to formulæ of $\mathcal{L}_{\mathtt{T}}$, likewise $\mathsf{PA}_{\mathtt{T}}$. If S is a classical theory we denote by S^i the appropriate theory based on intuitionistic logic. Let $\ulcorner . \urcorner$ denote some fixed primitive recursive Gödel coding of $\mathcal{L}_{\mathtt{T}}$. Though not essential for the present work, it is convenient to assume that $\ulcorner . \urcorner$ is total, i.e. every number is the code of some formula of $\mathcal{L}_{\mathtt{T}}$. Usual conventions apply to the arithmetisation of syntax that should be clear from the context in which they occur.

The base theory of truth we consider, called $\mathsf{Base}^i_{\mathtt{T}}$, extends $\mathsf{HA}_{\mathtt{T}}$ by the three axioms

1. $\forall x \forall y(\mathtt{T}x \wedge \mathtt{T}(x \dot{\rightarrow} y) \rightarrow \mathtt{T}y)$,
2. $\forall x(\mathtt{val}^i(x) \rightarrow \mathtt{T}x)$,
3. $\forall x(\mathtt{Ax}_{\mathsf{PRA}}(x) \rightarrow \mathtt{T}x)$.

Here $\mathtt{val}^i(x)$ expresses that x is the Gödel number of an intuitionistically valid first-order $\mathcal{L}_{\mathtt{T}}$-sentence and $\mathtt{Ax}_{\mathsf{PRA}}(x)$ expresses that x is the Gödel number of the universal closure of a non-logical axiom of primitive recursive arithmetic. The classical base theory of [4] and [8], called $\mathsf{Base}_{\mathtt{T}}$, is obtained from $\mathsf{Base}^i_{\mathtt{T}}$ by adding the law of excluded middle and the principles $\forall \ulcorner A \urcorner \mathtt{T}(\ulcorner A \vee \neg A \urcorner)$ stating that the underlying logic of the predicate $\mathtt{T}$ is classical. Thus $\mathsf{Base}^i_{\mathtt{T}}$ should be viewed as a natural base theory for the study of theories of truth over intuitionistic logic; it is a conservative extension of Heyting arithmetic and its only truth-theoretic commitments are that the collection of statements provably true is closed under modus ponens and comprises intuitionistic $\mathsf{I}\Sigma_1$.

Two theories are said to be *proof-theoretically equivalent* if they share the same theorems of $\mathcal{L}$ and, moreover, that this fact can be established within HA. We say a theory has *proof-theoretic ordinal* α if it is proof-theoretically equivalent to either $\mathsf{HA} + \mathtt{TI}(<\alpha)$ or $\mathsf{PA} + \mathtt{TI}(<\alpha)$ where $\mathtt{TI}(<\alpha)$ represents the schema of transfinite induction on initial segments of (a natural representation of) α.

Name	Axiom	Name	Axiom
Rep	$\mathtt{T}A \rightarrow \mathtt{TT}A$	Cons	$\neg(\mathtt{T}A \wedge \mathtt{T}\neg A)$
Del	$\mathtt{TT}A \rightarrow \mathtt{T}A$	Comp	$\mathtt{T}A \vee \mathtt{T}\neg A$
$\forall$-Inf	$\forall x \mathtt{T}A(\dot{x}) \rightarrow \mathtt{T}(\forall x A)$	$\vee$-Inf	$\mathtt{T}(A \vee B) \rightarrow \mathtt{T}A \vee \mathtt{T}B$
$\exists$-Inf	$\mathtt{T}(\exists x A) \rightarrow \exists x\, \mathtt{T}A(\dot{x})$	$\rightarrow$-Inf	$(\mathtt{T}A \rightarrow \mathtt{T}B) \rightarrow \mathtt{T}(A \rightarrow B)$

Name	Axiom Schema	Name	Rule of inference
In	$A \rightarrow \mathtt{T}A$	Intro	From A infer $\mathtt{T}A$
Out	$\mathtt{T}A \rightarrow A$	Elim	From $\mathtt{T}A$ infer A
		$\neg$Intro	From $\neg A$ infer $\neg\mathtt{T}A$
		$\neg$Elim	From $\neg\mathtt{T}A$ infer $\neg A$.

Table 1: Table of truth-theoretic principles

Table 1 shows the fifteen principles of truth considered in [4] and [7]. For presentation purposes these principles are given in shorthand and should be read in the usual more general formulation universally quantified with parameters (where applicable). For instance, the axiom schema $A \to \mathtt{T}A$ should be read as the schema $A(x_0, x_1, \ldots, x_n) \to \mathtt{T}\ulcorner A(\dot{x_0}, \dot{x_1}, \ldots, \dot{x_n})\urcorner$ for each formula A of $\mathcal{L}_\mathtt{T}$ with at most $x_0, \ldots, x_n$ free.

The following theorem outlines the known results regarding sets of principles over $\mathsf{Base}_\mathtt{T}$.

Theorem 1. Table 2 presents the complete list of maximal consistent sets of principles from table 1 over $\mathsf{Base}_\mathtt{T}$ with theories that are proof-theoretically equivalent alongside. The axioms Comp(w), $\vee$-Inf and $\to$-Inf are omitted from the list as they are all equivalent to Comp over $\mathsf{Base}_\mathtt{T}$.

Maximal consistent set	Equivalent theories
$\forall$. In, Intro, Rep, Del, Comp, $\neg$Elim, $\forall$-Inf, $\exists$-Inf.	PA
B. Rep, Cons, Comp, $\forall$-Inf, $\exists$-Inf.	ACA, $\mathsf{PA} + \mathtt{TI}(<\epsilon_{\epsilon_0})$
C. Del, Cons, Comp, $\forall$-Inf, $\exists$-Inf.	ACA, $\mathsf{PA} + \mathtt{TI}(<\epsilon_{\epsilon_0})$
D. Intro, Elim, Cons, Comp, $\neg$Intro, $\neg$Elim, $\forall$-Inf, $\exists$-Inf.	ACA_0^+, $\mathsf{PA} + \mathtt{TI}(<\varphi 20)$, $\mathsf{RA}_{<\omega}$
$\exists$. Intro, Elim, Del, Cons, $\neg$Intro, $\forall$-Inf.	Σ^1_1-DC_0, ID_1^*, $\mathsf{PA} + \mathtt{TI}(<\varphi\omega 0)$
F. Intro, Elim, Del, $\neg$Elim, $\forall$-Inf.	Σ^1_1-DC_0, ID_1^*, $\mathsf{PA} + \mathtt{TI}(<\varphi\omega 0)$
G. Intro, Elim, Rep, $\neg$Elim, $\forall$-Inf.	ACA_0^+, $\mathsf{PA} + \mathtt{TI}(<\varphi 20)$
H. Out, Elim, Del, Rep, Cons, $\neg$Intro, $\forall$-Inf.	ID_1, $\mathsf{KP}\omega$, $\mathsf{PA} + \mathtt{TI}(<\vartheta\epsilon_{\Omega+1})$
I. Rep, Del, Elim, $\neg$Elim, $\forall$-Inf.	ACA_0^+, $\mathsf{PA} + \mathtt{TI}(<\varphi 20)$

Table 2: Known results over classical logic

Isolation of the maximal consistent theories is due to Friedman and Sheard [4]. The proof-theoretic analysis of theories D and H is provided by Halbach [5] and Cantini [1] respectively. A lower bound on the strength of H is also present in [4]. The strength of the remaining seven theories is due to Leigh and Rathjen [8].

The consistency of sets of truth principles over $\mathsf{Base}^i_\mathtt{T}$ has also been explored. Leigh and Rathjen, in [7], classify all sets of truth principles as either consistent or inconsistent over $\mathsf{Base}^i_\mathtt{T}$. The present work completes the picture by determining the proof-theoretic strength of each resulting theory, outlined by the next theorem.

Theorem 2. Table 3 below lists all maximal consistent collections of the fifteen principles from table 1 over $\mathsf{Base}^i_\mathtt{T}$ together with proof-theoretically equivalent theories. For space considerations obvious redundancies in listing the axioms of each theory have been omitted.

At first glance it may not seem surprising that each of the theories $\forall^i$–I^i has the same proof-theoretic ordinal as its classical cousin. After all, the only axioms lacking compared to the classical theories is the law of excluded middle. The model construction used to verify consistency of the intuitionistic theories, however, have much in common with the fixed point construction utilised by Feferman in the analysis of KF, a theory whose intuitionistic formulation is a conservative extension of Heyting arithmetic. In the hope of providing a clearer understanding as to why, in contrast with KF, the theories do not collapse, we compare the ordinal analysis of F^i with that of F.

In [8] a classical sequent calculus $\mathsf{F}_\infty\frac{m}{\alpha}\Delta$ is defined into which F without Elim embeds and such that the following soundness result holds: If $\mathsf{F}_\infty\frac{m+1}{\alpha}\Delta$ is derivable and Δ consists of only formulæ in

Maximal consistent set	Equivalent theories
$\forall^i$. In, Intro, Rep, Del, Comp, ¬Elim, ∀-Inf, ∃-Inf.	HA
B^i. Rep, Cons, Comp, ∀-Inf, ∃-Inf.	ACA, $\mathsf{PA} + \mathtt{TI}(<\epsilon_{\epsilon_0})$
C^i. Del, Cons, Comp, ∀-Inf, ∃-Inf.	ACA, $\mathsf{PA} + \mathtt{TI}(<\epsilon_{\epsilon_0})$
D^i. Intro, Elim, Cons, Comp, ¬Intro, ¬Elim, ∀-Inf, ∃-Inf.	ACA_0^+, $\mathsf{PA} + \mathtt{TI}(<\varphi 20)$, $\mathsf{RA}_{<\omega}$
$\exists^i$. Intro, Elim, Del, Cons, ¬Intro, ∀-Inf, ∃-Inf, ∨-Inf.	Σ^1_1-DC^i_0, $\mathsf{HA} + \mathtt{TI}(<\varphi\omega 0)$
F^i. Intro, Elim, Del, ¬Elim, ∀-Inf, ∃-Inf, Comp(w), ∨-Inf.	Σ^1_1-DC^i_0, $\mathsf{HA} + \mathtt{TI}(<\varphi\omega 0)$
G^i. In, Elim, ∀-Inf, ∃-Inf, Comp(w), ∨-Inf, →-Inf.	ACA_0^{i+}, $\mathsf{HA} + \mathtt{TI}(<\varphi 20)$
H^i. Out, Rep, Cons, ∀-Inf, ∃-Inf, ∨-Inf.	ID^i_1, $\mathsf{KP}\omega^i$, $\mathsf{HA} + \mathtt{TI}(<\vartheta\epsilon_{\Omega+1})$
I^i. Rep, Del, Elim, ¬Elim, ∀-Inf, ∃-Inf, Comp(w), ∨-Inf.	ACA_0^{i+}, $\mathsf{HA} + \mathtt{TI}(<\varphi 20)$

Table 3: New results over intuitionistic logic

which the truth predicate occurs positively then some element of Δ becomes a true sentence if the truth predicate is interpreted as the set of sentences B for which $\mathsf{F}_\infty \vdash^{m}_{\varphi(m+1)\alpha} B$ is derivable, where φ represents the two-placed Veblen function. It is crucial for the soundness result above that the calculus enjoys cut elimination, since then if $\mathsf{F}_\infty \vdash^{m}_{\alpha} \mathtt{T}s$ is derivable, there is a derivation in which any sub-derivation with the same rank m has the form $\mathsf{F}_\infty \vdash^{m}_{\gamma} \mathtt{T}s_1, \mathtt{T}s_2, \dots, \mathtt{T}s_n$. It follows that the consistency proof for F can be formalised in ID^*_1, the sub-theory of ID_1 in which proof by induction and the schema of induction is only stated for formulæ that do not contain negative instances of other fixed point predicates. This is because the properties required to establish that F is interpretable in the calculus $\mathsf{F}_\infty \vdash$, specifically cut elimination and closure under Elim, are purely positive in their expression, being of the form "if X is derivable in the sequent calculus, then so is Y". It is known, however, that F cannot be embedded into ID^*_1 by interpreting the truth predicate simply as membership of some positive inductive definition because the latter theory does not contain the full schema of induction, whereas the former theory does. One would expect the study of F to offer some explanation as to why this is the case, but it is surprisingly silent on the issue. Indeed it is only from the analysis of F^i, which is more involved, that we begin to find some answers.

As with the analysis of F, the first step towards achieving an upper bound for F^i is to embed the theory without the rule Elim into an infinitary sequent calculus. We thus define $\Gamma \Rightarrow^m_\alpha A$ according to the usual rules for an infinitary version of HA together with rules representing each axiom of F^i and the rule Intro, so that if A is a sentence derivable in F^i without using Elim and with at most m applications of Intro then $\emptyset \Rightarrow^m_\omega A$. To embed all theorems of F^i one needs to establish that the calculus admits the rule Elim. The rule cannot be added directly to $\Rightarrow$ because of its impact on cut elimination, a key factor in obtaining the final interpretation into arithmetic.

The second step is to prove a soundness result similar to the one mentioned above. In this case we cannot make do with analysing only $\mathtt{T}$-positive derivations. The rules corresponding to Comp(w) and the axiom $\mathtt{T}(A \to B) \wedge \mathtt{T}A \to \mathtt{T}B$ require negative occurrences of the truth predicate in their premises, so unlike with F, it is not obvious that $\Rightarrow$ permits the elimination of cuts. As such, to deduce the admissibility of Elim in the context of F^i, a more general soundness result is required, one that applies to derivations involving sequents with both positive and negative occurrences of the truth predicate. The solution is to make use of asymmetric interpretations for truth to obtain an analogous soundness result. These are formalised models in which positive and negative occurrences of the truth predicate may have different interpretations. In this setting they take the following form.

Lemma. There is a hierarchy of intuitionistic Kripke models $(\mathfrak{F}^m_\alpha \mid \alpha < \Gamma_0, m < \omega)$ and normal functions $\{f^m_\alpha \mid m < \omega\}$ such that i) if $\mathfrak{F}^m_\alpha \models \mathtt{T}\ulcorner A \urcorner$ then $\emptyset \Rightarrow^n_{f^m_\alpha(0)} A$ is derivable for some $n < m$; and ii) if $\Gamma \Rightarrow^m_\alpha A$ is derivable and $\Gamma \cup \{A\}$ consists solely of atomic formulæ, then whenever $\mathfrak{F}^m_\gamma$ satisfies all elements in Γ, $\mathfrak{F}^m_{f^m_\alpha(\gamma)}$ satisfies A.

For f^m_α we pick the function $\gamma \mapsto \varphi m(\gamma + \omega^\alpha)$; the model $\mathfrak{F}^m_\alpha$ is then defined to be a linear Kripke ω-structure with two worlds. In the top world every sentence is considered true, while the second world is defined so that (i) in the above lemma holds. Therefore the rule Elim is admissible in $\Rightarrow$ and if $\mathsf{F}^i \vdash A$ then the universal closure of A is derivable with finite m and height bounded by $\varphi\omega 0$. In order to deduce that F^i is interpretable in $\mathsf{HA} + \mathtt{TI}(<\varphi\omega 0)$ it is necessary to establish that the resources required to prove lemma do not go beyond what is expressible in HA itself. Thus reference to the models $\mathfrak{F}^m_\alpha$ must be removed in favour of arithmetically definable notions. This is possible because $\Rightarrow$ enjoys partial cut elimination, so any derivation of $\Gamma \Rightarrow^m_\alpha A$ satisfying the assumption of (ii) in the lemma has a derivation involving atomic formulæ only.

The use of asymmetric interpretations appears essential. For if an interpretation is available in which the truth predicate is directly replaced by a notion of derivability then it would be expected that F^i would be interpretable in the intuitionistic version of ID^*_1. Since the latter theory is a conservative extension of arithmetic this is impossible. One direction for future research is to determine an interpretation of F into ID^*_1. We expect more success in this matter by looking at variants of the above asymmetric interpretation rather than at direct interpretations motivated by the classical analysis.

A classical theory with proof-theoretic ordinal $\varphi\omega 0$ that retains its strength when transferred to intuitionistic logic is the theory of Σ^1_1 dependent choice (Σ^1_1-DC_0). It is reasonable to suppose there is a correlation between this axiom schema and the theories F and F^i. Thus another avenue of interest is to explore further connections between axiomatisations of truth and formulations of choice in second-order analysis.

Bibliography

[1] Cantini, A. A Theory of Formal Truth Arithmetically Equivalent to ID1. *Journal of Symbolic Logic 55*, 1 (1990), 244–259.

[2] Feferman, S. Toward useful type-free theories. I. *The Journal of Symbolic Logic 49*, 1 (1984), 75–111.

[3] Field, H. *Saving truth from paradox*. Oxford University Press, 2008.

[4] Friedman, H., and Sheard, M. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic 33* (1987), 1–21.

[5] Halbach, V. A System of Complete and Consistent Truth. *Notre Dame Journal of Formal Logic 35*, 3 (1994), 311–327.

[6] Halbach, V., and Horsten, L. Axiomatizing Kripke's theory of truth. *Journal of Symbolic Logic 71*, 2 (2006), 677–712.

[7] Leigh, G. E., and Rathjen, M. The Friedman-Sheard programme in intuitionistic logic. *Journal of Symbolic Logic 77*, 3 (2012), 777–806.

[8] Leigh, G. E., and Rathjen, M. An ordinal analysis for theories of self-referential truth. *Archive for Mathematical Logic 49*, 2 (Jan. 2010), 213–247.

Revision without ordinals

Edoardo Rivello
Scuola Normale Superiore, Pisa

Roughly speaking, semantic theories of Truth deal with the problem of constructing a model for a given language together with an interpretation of the Truth-predicate for this language. To set things in a more suitable way for a mathematical treatment we can assume — as it is often the case in the literature on this topic — that (1) a formalized first-order version of Peano Arithmetic (PA) plays the role of the object language, (2) a self-referential Truth-predicate T is added to the language $\mathcal{L}$ of PA and applies to sentences of the extended language $\mathcal{L} + \{\mathsf{T}\}$ and (3) the interpretation of T over the standard model of PA has to be built up within Zermelo-Fraenkel (ZF) set theory (or within some weaker fragment of ZF).

Several semantic theories of self-referential Truth can be presented as systems of ordinal-length sequences of attempts to give an interpretation for the Truth-predicate. This strategy provides a general mathematical framework in which different approaches to Truth can be fruitfully compared: for an analysis of this kind see Barba [1] where both constructions based on *approximation sequences* (as in Kripke [6]) and constructions based on *revision sequences* (Gupta and Belnap [4] is the standard reference) are covered.

While some knowledge of natural numbers seems to be unavoidable in inquiring about Truth — since, at least, we need a theory of the syntax of the truth-bearers (sentences, in our setting) and this theory comes to be equivalent to some weak form of arithmetic — the introduction of heavy set-theoretical assumptions in the metatheory could be regarded as an unwelcome accident in a theory of Truth.

As a matter of fact, the approximation constructions can also be presented in an inductive style, without any reference to the ordinals, so reducing the amount of set-theoretical notions needed in the metalanguage (see Fitting [3] for a detailed treatment). Apparently, this possibility constitutes a substantive difference between the approximation and the revision approaches: but it will be shown that this is not the case.

From a result by McGee [8] we already know that, in inquiries about Truth for countable languages, we can restrict ourselves to countable revision sequences. This fact suggests that it might be possible to do the Revision Theory without ordinals just by reducing ordinals to natural numbers via some index notation for countable ordinals.

I will present here a different approach to this problem that directly shows that the mathematical features of revision sequences that are relevant for the Revision Theory of Truth can be implemented in an ordinal-free setting — worked out in the Dedekind-Kuratowski tradition — as in the approximation case.

This strategy of doing revision without ordinals can show its merits under several aspects:

- General motivations for eliminating the ordinals from some pieces of mathematics can be found in Kuratowski [7]:

  > "Even though, sometimes, transfinite numbers [ordinals] can be shown to be fruitful in making the exposition shorter or easier, the existence of a process that allows

to avoid ordinals, in proving theorems that do not deal with the transfinite, is important for the following two reasons: in reasoning about ordinals we implicitly appeal to axioms that ensure their existence; but to weak the axioms system that we use in proving something is desirable both from a logical and from a mathematical point of view. Moreover, this strategy expunges from the arguments the unnecessary elements, increasing their aesthetic value.[4]"

- Our ordinal-free presentation of the Revision Theory of Truth, together with Fitting's ordinal-free exposition of Kripke's theory, provides a framework in which to compare approximation and revision approaches to Truth that constitutes an alternative to Barba's analysis based on ordinal-length sequences.
- An ordinal-free presentation makes easier to evaluate how the Revision Theory of Truth is sensitive to its underlying set theory: both in order to minimise the set-theoretical assumptions taken from the Zermelo-Fraenkel axiomatization and in order to explore the possibility of doing revision also in alternative set theories.
- Avoiding ordinals may help in facing the problem of lifting the revision-theoretical approach from "toy" object languages like arithmetic to more complex languages, since the entire set-theoretical process of revision can be recast in higher-order logical terms.
- The ordinal-free approach to the Revision Theory of Truth might be shown valuable also from a heuristic point of view, helping to focus our attention on the intrinsic circularities exhibited by the concept of Truth rather than on the redundancies introduced in the analysis by the transfinite iteration of the revision operator.

I plan to illustrate some of the above points as follows. First, I will briefly recall an overview of the Revision Theory of Truth. Then, I will introduce the mathematical notion of *generalized orbit* and its basic properties. Finally, we discuss how this latter concept can be successful in replacing the notion of *revision sequence* as the fundamental notion for a theory of Truth.

Bibliography

[1] J. Barba: *Construction of Truth Predicates: Approximation vs. Revision*, The Bulletin of Symbolic Logic 4 n. 4: 399-417, 1998

[2] R. Dedekind: *Was sind und was sollen di Zahlen?*, 1888

[3] M. Fitting: *Notes on the Mathematical Aspects of Kripke's Theory of Truth*, Notre Dame Journal of Formal Logic 27 n. 1: 75-88, 1986

[4] A. Gupta, N. Belnap: *The revision theory of truth*, 1993

[4]"Bien que l'emploi des nombres transfinis puisse présenter quelquefois quelques avantages au point de vue de la brièveté et de la simplicité, l'existence d'un procédé permettant de supprimer la notion de ces nombres dans les démonstrations des théorèmes qui ne concernent guère le transfini est importante pour les deux raisons suivantes: en raisonnant avec le nombres transfinis on fait implicitement usage de l'axiome de leur existence; or, la réduction du système d'axiomes employés dans les démonstrations est désirable au point de vue logique et mathématiques. En outre, cette réduction affranchit les raisonnements de l'élément qui leur est étranger, ce qui augment leur valeur esthétique" (the English translation is mine).

[5] H. G. Herzberger: *Notes on Naive Semantics*, Journal of Philosophical Logic 11 n. 1: 61-102, 1982

[6] S. Kripke: *Outline of a theory of truth*, Journal of Philosophy 72: 690-716, 1975

[7] C. Kuratowski: *Une méthode d'élimination de nombres transfinis des raisonnements mathémathiques*, Fundamenta Mathematicae 3: 76-108, 1922

[8] V. McGee: *Truth, Vagueness and Paradox. An essay on the logic of truth*, 1991

The disentanglement of syntax from a model theoretic point of view

Andrea Strollo
University of Helsinki

In the field of formal theories of truth a prominent place is occupied by the (broadly) Tarskian theory. The theory is also called "there is a (full-not inductive) satisfaction class" or, shortly, PA$(S)^-$. It consists, apart from the axioms of the base theory PA, of the truth-compositional axioms (here I consider arithmetical induction only):

1. $\forall\varphi(\text{Atomic}(\varphi) \rightarrow (\text{Tr}(\varphi) \leftrightarrow \text{Tr}^*(\varphi)$;
2. $\forall\varphi(\text{Tr}(\neg\varphi) \leftrightarrow \neg\text{Tr}(\varphi)$;
3. $\forall\varphi\forall\psi(\text{Tr}(\varphi \wedge \psi) \leftrightarrow \text{Tr}(\varphi) \wedge \text{Tr}(\psi))$;
4. $\forall\varphi\forall i(\text{Tr}(\forall v_i\varphi) \leftrightarrow \forall t\text{Tr}(\varphi(t/v_i))$;[5]

The name PA$(S)^-$ should be explained, because it carries important information and leads to debatable aspects.

When studying truth theories, it is often said that a background theory of syntax is needed. Without it, writing down axioms for a truth predicate and working out simple operations is impossible. We intend to ascribe truth to so called "truth-bearers"; a theory of syntax is intended to give us basic information about how these entities behave. One would expect a theory of syntax to consist of principles about linguistic expressions, and this was exactly the case in the original work of Tarski. However, explicit formal theories of syntax, in the style of concatenation theories i.e., are not much widespread among truth-theorists.

The reason is that, after Gödel, we know that a very good deal of syntax can be developed inside PA (as in even weaker arithmetical theories). According to Gödelization, we can correlate natural numbers and symbols of the language of PA. There are many ways to think of this correspondence between strings of symbols and numbers, but one often adopted is the easiest one: strings of symbols are identified with corresponding numbers.

Since PA is a very well known theory and we often want to consider the effects of adding a theory of truth to a theory of arithmetic like PA, it is clear that PA is the best candidate for a theory of syntax in many cases. Indeed, we can add our theory of truth directly to PA, without describing an independent theory of syntax. All of this sounds completely familiar and unproblematic, but it happens that a lot of intricacies survive here.

[5] This way of writing is comfortable but incorrect. A lot of coding apparatus has been suppressed to achieve a greater readability. To be rigorous we should write axiom 2, for example, like this: $\forall x\forall y(\text{Sent}(x) \wedge \text{Sent}(y) \wedge \text{Neg}(y,x) \rightarrow (\text{Tr}(y) \leftrightarrow \neg\text{Tr}(x)))$. Here I shall persist with the most perspicuous presentation, but keep in mind that this is the right form.

Among the many syntactical properties that can be represented in PA, we can obviously define that of being a sentence of the language of PA. For example there is a formula 'Sent(n)' which is true of n if and only if n is the code of a sentence of the language of PA. Until we consider the standard model N, as it is natural doing, this works as expected. PA, however, also has non-standard models. In those non-standard models, because of Overspill, the formula 'Sent(x)' is going to be satisfied by non-standard numbers too. Namely, if $M \models$ PA is non-standard, we have that $M \models$ Sent(b) for some $b \in M$, and b non-standard. The existence of non-standard numbers that, according to the model, code "sentences", drags us towards the realm of non-standard sentences. Very roughly, non-standard sentences are sentences with a "non-standard structure".[6]

Since every axiom of PA$(S)^-$ is subjected to a clause stating that the truth predicate applies to elements satisfying the formula 'Sent(x)', when we have a non-standard model, non-standard numbers can well enter into the range of the truth predicate. Actually this is not only possible but mandatory. In fact, PA$(S)^-$ proves $\forall\varphi\{\text{Sent}(\varphi) \to [\text{Tr}(\varphi) \vee \neg\text{Tr}(\varphi)]\}$, thus, for every φ such that $M \models \text{Sent}([\varphi])$ either φ or $\neg\varphi$ must be in the extension of "Tr", even if φ is non-standard.

Now let S be this extension. Namely, given a model $M \models$ PA, S is the set of numbers satisfying the axioms of PA$(S)^-$, $M, S \models$ PA$(S)^-$. When a model M has such a set S, we say that S is a satisfaction class for M. This explains the name PA$(S)^-$.

One of the most notable results about PA$(S)^-$ is the Lachlan's theorem, according to which not every model of PA has a satisfaction class. This result becomes even more surprising confronted with the fact that PA$(S)^-$ is (proof theoretically) conservative over PA. Thus, that some models are excluded is not due to new theorems in the language of PA.

Because of conservativity, one may well find — or, at least, I really found it so — the fact that in some model M we cannot have a suitable extension for the truth predicate to be a puzzling result. Actually, it really seems that something, somewhere, is gone wrong. Ingenuously, it seems that a set S of sentences satisfying the axioms of PA$(S)^-$ is always available, actually. Aren't we entitled to do model theory in every model of PA? After all, a set of sentences satisfying PA$(S)^-$ exists in every M, as "$M \models$" proves. But, according to Lachlan's theorem, we cannot exploit this fact. What is crucial is the existence, in non-standard models, of non-standard sentences, which make Lachlan's proof work. The problem is not only technical, it is a general philosophical one, since we do not recognize non-standard sentences as real sentences and we shouldn't be forced to apply our theory of truth to them, despite their mathematically interesting nature.

The source of the phenomenon is not hard to identify. It lays in the fact that, when writing down the Tarskian axioms for truth, we quantify in the object language, the language of PA, through the formula 'Sent(x)'. Doing so, anytime we have a non-standard model M of the base theory, we are forced to consider non-standard elements satisfying it, obtaining non-standard sentences. This is symptomatic of the fact that we developed our syntactic theory in the base theory. We have one single theory doing two different jobs. On the one side we treat it as our base arithmetical theory, on the other side as our syntactical theory. Having one single theory simplifies the issue, but forces the two functions to overlap and intertwine. This gives unpleasant results: in one precise sense, the base language has no non-standard sentences, and we would like to have a theory of syntax[7] not forced to them only because the base theory has non-standard models. In the present framework, instead, non-standard numbers

[6] For example, a sentence with a non-standard number of conjuncts

[7] Clearly, a theory of syntax may also have non standard models. The problem, though, is that here the models of the base theory and those of the syntax theory are the same models, so that what happens at the level of arithmetical objects, also happens at level of syntactical objects.

automatically become non-standard sentences. If the base theory is about non-standard objects, we must conclude that we have non-standard sentences in the object language.

Clearly, we are not forced to take seriously non-standard sentences when we specify the semantics for the language of PA in general. Non-standard sentences are a side-effect of our formalization. Their existence shows that such a formalization makes incomplete justice to our meta-mathematical reasoning about the semantics of the language of PA.

The natural solution to avoid these kinds of side-effects is that of restoring the original Tarskian attitude, namely distinguishing neatly the theory of syntax from the base theory. For sake of brevity, here I give just a quick sketch to get a rough idea of how we can proceed. We start with a theory of syntax S in the language L_S.[8] Since we want to devise syntax and semantics for a further language, say L_{PA}, we also need to be able to refer to the expressions of L_{PA}, and the objects L_{PA} is about. The simplest way is then to adopt the language $L_S \cup L_{\text{PA}}$. Some precaution should be certainly adopted here, in order to take the intended domains apart. At this point we can add semantic axioms, namely axioms for truth and denotation.[9] Call the complex theory of truth (for L_{PA}) we obtain "Σ".

It can be proved that every model M of PA can be expanded to a model of Σ. The exact proof will depend on the technical details, that, of course, must be spelled out; what is important, however, is the simple idea of this proof. According to it, this time, we can restrict our attention to standard sentences: we can interpret, by simple stipulation, the expressions of the object language in the standard part of M. Disentangling the theory of syntax from the base theory, then, gives us what we wanted: we are not forced to consider non-standard sentences in our truth theory, just because the base theory has non-standard models. Being able to deal with standard sentences only allows us to circumvent Lachlan's theorem, since a set S, satisfying compositional axioms, is available in any model of PA; in particular, it is a set of standard sentences, as expected.

The idea of a disentanglement of the theory of syntax from the base theory has been (re)proposed very recently by Richard Heck, who argued from mere proof-theoretic considerations. Basically, he stressed that, using a truth theory, we can, unpleasantly, prove consistency statements for PA already in weaker fragments of PA. This can be done - cutting a very and interesting story short - only because we go freely back and forth from syntax to numbers, inside the same theory. If we disentangle syntax, it becomes clear that those consistency statements belong to the theory of syntax only, and the unfortunate results are overcome.

Volker Halbach stressed that there is something puzzling in this approach as well. As a matter of fact we know that the truth of consistency statements, formulated in a metalanguage, corresponds to the truth of certain statements in the object language. This aspect is lost in Heck's framework. If we want to make justice to our meta-mathematical reasoning and formalize it adequately, then, a formal reconstruction of Gödelization, and bridge laws connecting the two sides, must be added.

I see the simple model theoretic considerations above as a different way to arrive at the same attitude of Heck and Halbach. In particular, I take these reflections as another good reason for disentangling syntax and, at the same time, as a confirm of Halbach's diagnosis: the identification between numbers and expressions is certainly legitimate and mathematically highly useful, but it should handled with extraordinary care.

[8] We may also keep using PA as our syntax theory, provided that we label and use it as a different theory from the arithmetical base theory.

[9] The addition of separate axioms for denotation, or for sequences in case we have a satisfaction predicate, is a little tricky. Notice however, that no trick is needed until we do not disentangle the syntax from the base theory. In particular, we have variables, which are syntactical objects, already identified with numbers, so a bridge theory connecting the two domains is not needed.

Bibliography

[1] Craig, W.; Robert V. Finite axiomatizability using additional predicates. *Journal of Symbolic Logic 23*, (1958), 289–308.

[2] Engström F. *Satisfaction Classes in Non standard Models of First-order Arithmetic,* Thesis for the Degree of Licentiate of Philosophy, Department of Mathematics Chalmers University of Technology and Göteborg University, 2002.

[3] Halbach, V. *Axiomatic Theories of Truth,* Cambridge University Press, Cambridge, 2011.

[4] Heck, R. *The strength of Truth Theories,* draft, 2009.

[5] Kaye, R. *Models of Peano Arithmetic,* Oxford Logic Guides, Oxford University Press, 1991.

Yablo's paradox and ω-inconsistency

Shunsuke Yatabe
Graduate School of Letters, Kyoto University

We focus on the *co-inductive* character of Yablo's paradox, analyzing it by comparison in truth theories and in **ZFA**. We show that the ω-inconsistency of truth theories is because, while they allow mixtures of induction and co-induction, such mixtures are impossible in an ω-consistent **ZFA**.

Introduction

Let us assume there exist infinitely many propositions $\langle S_0, S_1, S_2, \cdots \rangle$ such that S_n insists that S_i is false for any $i > n$. Then these propositions imply a contradiction. First let us assume S_0 is false. Then there must be $j > 0$ such that S_j is true. This means that all $S_{j+1}, S_{j+2}, S_{j+3}, \cdots, S_k, \cdots$ must be false. However, if S_{j+1} is false, then there exists $k > j+1$ such that S_k is true, a contradiction. Next assume S_0 is true. Then $S_1, S_2, \cdots$ are false, identical to the previous case. This is the well-known Yablo's paradox [Yab93].

There has been previous discussion as to whether Yablo's paradox is self-referential, but this paper does not address that topic. Instead, we focus on how Yablo propositions are constructed. The answer to whether Yablo's paradox is self-referential seems to depend on how these propositions are constructed, and the essence of their construction. A source of trouble is that only one method of constructing $\langle S_n : n \in \omega \rangle$ is known, using diagonalization in a truth theory in classical logic [P97]. Such deficiency of comparison examples could lead to a mistake that we regard properties that contingently hold in the truth theory (and do not hold in other theories) as essential properties of Yablo's paradox. For example, consistent truth theories with sufficient expressive power to define Yablo propositions should be ω-inconsistent [L01], but we do not know whether ω-inconsistency is essential in Yablo's paradox.

Yablo propositions satisfy a characteristic property in that the intuitive meaning of S_i is $\bigwedge_{j>i} \neg \mathbf{Tr}(\ulcorner S_j \urcorner)$ (if the language has an infinite conjunction). Therefore,

$$\begin{array}{rcl} S_i & \equiv & \neg \mathbf{Tr}(\ulcorner S_{i+1} \urcorner) \wedge S_{i+2} \\ \neg S_i & \equiv & \mathbf{Tr}(\ulcorner S_{i+1} \urcorner) \vee \neg S_{i+2}. \end{array}$$

This means each S_i is constructed by directly using S_{i+1} and S_{i+2}. However, to construct S_{i+2}, we need S_{i+3} and S_{i+4}, etc. In this way, there is an infinite regress; we need infinitely many $\langle S_{i+1}, S_{i+2}, S_{i+3}, \cdots \rangle$ to construct S_i in the end. The characteristic points of this construction are that (1) we only directly use finitely many already-constructed objects to construct a new object, and (2) we need infinitely many steps to reach the *initial* construction case (meaning this is not inductive construction).

Such constructions are called *co-inductive*, and are widely used in computer science to represent behaviors of non-terminate automatons [C93] because they allow construction of *potentially* infinite objects in a finite way. Yablo's paradox seems to be evidence that co-induction is naturally used in natural language. We focus on the *co-inductive* character of Yablo's paradox, and analyze it by comparing the paradox in truth theories to that in **ZFA**.

Preliminaries on ZFA

One of the most famous ways to define a co-inductive language, a language with co-inductively defined formulae, is to use **ZFA** [BE87] [BM96]. This is done by coding co-inductively defined propositions by hypersets. As for Yablo's paradox, Yablo suggested fixing **ZFA** as an analysis framework [Yab06], but abandoned this approach without serious consideration. **ZFA** is an axiomatic set theory, **ZF** minus the axiom of foundation plus the anti-foundation axiom (**AFA**), which allows definition of *hypersets*, which need not be well founded in classical logic. Due to space limitations, we present the so-called *flat system lemma* with only a brief review.

Definition 1. A flat system of equations $\langle X, A, e\rangle$ has the following characteristics:

- $X \subseteq U$ (urelements, interpreted as variables),
- A is an arbitrary set, and
- $e : X \to \mathcal{P}(X \cup A)$.

An example of a flat system is $\langle\{a\}, \emptyset, \{\langle a, \{a\}\rangle\}\rangle$ for some urelement a; since $e(a) = \{a\}$, this system represents an equation $x = \{x\}$, where x is a free variable.

Theorem 3. **AFA** guarantees that *any flat system of equations defines hypersets uniquely.*

As a sort of co-inductive definition[10], consider the flat system

$$\langle\{a_n : n \in \omega\}, \emptyset, \{\langle a_n, \{a_{n+1}, a_{n+2}\}\rangle : n \in \omega\}\rangle,$$

which represents equations $x_n = \{x_{n+1}, x_{n+2}\}$ for any n (the construction is finite in any successor step but we cannot achieve this in the initial case).

We fix **ZFA** as the framework of this paper because, thanks to [BE87], it is one of the most famous truth theory frameworks that enables purely co-inductive construction of formulae[11]. The framework of [BE87] seems to be *overkill* for semantic paradoxes. The liar proposition can be represented even as arithmetic, but **ZFA** produces hypersets, as many as ordinal well-founded sets, to represent such paradoxical propositions. The real value of this framework is that it allows many kinds of co-inductive construction[12].

---

[10] Actually **ZFA** is a set theory whose sets are constructed by co-induction in some transfinite induction step. The universe of **ZFA** is constructed by $\mathbf{V}_0 = \emptyset$, $\mathbf{V}_{\alpha+1} = \mathbf{V}_\alpha \bigcup \mathcal{P}^*(\mathbf{V}_\alpha)$ and $\mathbf{V}_\gamma = \bigcup_{\delta<\gamma} \mathbf{V}_\delta$ for any γ limit, where $\mathcal{P}^*(A) = \{x :\in |_{\mathrm{TC}(x)}$ is bisimilar to R for some $R \subseteq \mathbf{TC}(A)^2\}$[V04].

[11] Many theories allow co-inductive object definitions. For example, an intuitionistic theory has been extended to allow such definitions (we do not have to worry about overly rich ontologies in such theories) [C93], and naive set theories in non-classical logics have strong co-inductive characters [Yat12a]. However, **ZFA** is the most well known among them.

[12] As Yablo pointed out in [Yab06], there is a counterintuitive problem that any propositions S_i, S_j of Yablo's paradox are mutually identical. If we fix an Austinian-like approach, all propositions are pairwise distinct (*situations* are taken into consideration). We omit the details here due to space limitations.

Coding Yablo propositions by hypersets

Let us introduce the construction of *Russellian propositions* or *Austinian types*[13]. Define their *co-inductive* coding method by hypersets as follows:

Definition 2 (Russellian propositions or Austinian types). Formulae are coinductively coded in **ZFA** as follows:

- $\lceil A \wedge B \rceil = \{\{\mathbf{c}, \lceil A \rceil\}, \{\mathbf{c}, \lceil B \rceil\}\}$ and $\lceil \wedge_{i \in I} A_i \rceil = \{\{\mathbf{c}, \lceil A_i \rceil\} : i \in I\}$,
- $\lceil A \vee B \rceil = \{\{\mathbf{d}, \lceil A \rceil\}, \{\mathbf{d}, \lceil B \rceil\}\}$ and $\lceil \vee_{i \in I} A_i \rceil = \{\{\mathbf{d}, \lceil A_i \rceil\} : i \in I\}$,
- $\lceil \neg A \rceil = \{\mathbf{n}, \lceil A \rceil\}$,
- $\lceil \mathbf{Tr}(A) \rceil = \{\mathbf{t}, \lceil A \rceil\}$

for some fixed set $\mathbf{c}, \mathbf{d}, \mathbf{n}, \mathbf{t}$ which are not equal to any natural numbers.

Note that this coding does not have an initial case, but is sufficient to code the liar propositions or Yablo propositions. For example, the liar proposition λ is coded by a Russellian proposition $\lceil \lambda \rceil$ satisfying $x = \{0, \{*, x\}\}$.

Next let us define Yablo propositions[14].

Definition 3 (Yablo propositions). Yablo (Russellian) propositions $\{S_n : n \in \omega\}$ are coded by the following equation: let $\langle \{x_n, p_n : n \in \omega\}, \{\mathbf{c}, \mathbf{n}, \mathbf{t}\}, e \rangle$ be an infinite flat system such that, for any $n \in \omega$,

$$
\begin{aligned}
e(x_n) &= \{p_k : k > n\} \\
e(p_n) &= \{\mathbf{c}, q_n\} \\
e(q_n) &= \{\mathbf{n}, r_n\} \\
e(r_n) &= \{\mathbf{t}, x_k\}
\end{aligned}
$$

Then $S_0, S_1, \cdots$ are solutions of $x_0, x_1, \cdots$.

Theorem 4. Yablo (Russellian) propositions $\langle S_n : n \in \omega \rangle$ exists in **ZFA**[15].

[13] Roughly speaking, an Austinian proposition is a pair of a *situation* and an Austinian type: different definitions of situations give different definitions of propositions.

[14] We do not need the truth predicate to construct Yablo propositions in this framework. $\langle Y_n : n \in \omega \rangle$ are defined by $Y_0, Y_1, \cdots$ are solutions of $x_0, x_1, \cdots$ and $\neg Y_0, \neg Y_1, \cdots$ are solutions of $y_0, y_1, \cdots$ appearing in

$$
\begin{aligned}
e(p_n) &= \{\mathbf{c}, y_n\} \\
e(q_n) &= \{\mathbf{c}, x_n\} \\
e(x_n) &= \{p_k : k > n\} \\
e(y_n) &= \{q_k : k > n\}
\end{aligned}
$$

The intuitive meaning of Y_n is $\wedge_{n<i} \neg Y_i$, and this is equivalent to $\neg Y_{i+1} \wedge Y_{i+2}$. Recall that the liar paradox is not unique but an instance of a self-referential paradox; a Russell paradox is another. In this sense, Yablo's paradox is just an instance of a *co-inductive* paradox.

[15] Yablo's paradox implies a contradiction when applying Russellian semantics. If we apply Austinian-like semantics, all Yablo propositions are simply false (and thus do not imply a contradiction)

The proof is a simple application of theorem 3. We note that, as we pointed out, $S_i = S_j$ holds for any i, j since there is a *bisimulation* among all $\in$-graphs $\langle S_n : n \in \omega \rangle$ by this coding[16][17]. However, this is just a technical problem: just adding indexes makes them mutually different hypersets [Yat12b] (but we omit the detail because they are essentially the same). We also note that any S_n forms an infinite-branching tree of infinite height[18].

A comparison to truth theories: a source of ω-consistency

As discussed above, well-known consistent theories with sufficient expressive power, like Γ [Mc85] and $\mathbf{CT}_\omega$ [HH05], are ω-inconsistent. Yablo propositions $\langle \bar{S}_x : x \in \omega \rangle$ are constructed by the fixed point lemma in such theories as follows:

$$\bar{S}_x \quad \equiv \quad (\forall z)[z > x \to \neg \mathbf{Sat}(\ulcorner S_x \urcorner, z)],$$

where $\mathbf{Sat}(\ulcorner \varphi(x) \urcorner, z) \equiv \mathbf{Tr}(\ulcorner \varphi(z) \urcorner)$. Roughly speaking, the intuitive meaning of $\bar{S}_x$ is

$$\bar{S}_x \equiv \underbrace{\cdots \wedge \neg \mathbf{Sat}(\ulcorner \bar{S}_x \urcorner, x+2) \wedge \neg \mathbf{Sat}(\ulcorner \bar{S}_x \urcorner, x+1)}_{\infty \text{ many}}$$

The main difference between this construction and that of **ZFA** is whether the construction has an *initial* case or not. In the **ZFA** case, the construction does not have an initial case. Truth theory constructions do have an initial case S_x, however, and any S_y is constructed from S_x as $\mathbf{Sat}(\ulcorner S_x \urcorner, y)$ for any $y > x$. Thanks to the truth predicate, the fixed point lemma enables an infinite operation over formulae (S_x itself is a limit of infinite operation $\wedge_{y>x} \neg \mathbf{Sat}(\ulcorner S_x \urcorner, y)$). The construction of S_x is not by pure co-induction, but by *a mixture of induction and co-induction*, that is, a co-inductive construction with the initial case.

This mixture plays a key role in the proof of ω-inconsistency in truth theories. For example, in Γ [Mc85], ω-inconsistency is proved by the following formula γ:

$$\begin{aligned} \gamma \quad &\equiv \quad \neg \forall x \mathbf{Tr}(f(x, \ulcorner \gamma \urcorner)) \\ f(n, \ulcorner \varphi \urcorner) \quad &= \quad \ulcorner \underbrace{\mathbf{Tr}(\ulcorner \cdots \mathbf{Tr}}_{n \text{ times}}(\ulcorner \varphi \urcorner) \cdots) \urcorner \end{aligned}$$

Roughly speaking, γ is defined by *a mixture of induction and co-induction* in the sense that the intuitive meaning of γ is $\gamma \equiv \neg \underbrace{\mathbf{Tr}(\ulcorner \cdots (\ulcorner \mathbf{Tr}}_{\infty \text{ many}}(\ulcorner \gamma \urcorner)\urcorner) \cdots)\urcorner)$.

[16] Let us consider the meaning of this. In the paradox, first we take S_0 and assume it is true (or false). However, even though we first assume S_i is true (false), the *behavior* of the paradox, the derivation of the inconsistency, is an identical form. If we formalize the paradox using game semantics, the player who gives a counterexample has a very simple winning strategy regardless of the opponent's choice, S_0 or S_i. In this sense, S_0 and S_i are identical. Of course, the difference in the starting point can be distinguished if we consider the hidden parameter, *situations*: we can distinguish S_0 and S_i in Austinian-like Semantics.

[17] Note that the mutual equality of Yablo propositions collapses Yablo's paradox to a simple liar-like self-referential paradox. Actually, since $S_0 = S_i = S$, the paradox, $S_0 \to \neg S_i \wedge S_i$ and $\neg S_0 \to S_i \wedge \neg S_i$, are just equal to $S \to \neg S$ and $\neg S \to S$.

[18] Each tree S_n is self-similar, i.e., for any branch t of S_n, there is a sub-tree $T \subseteq S_n|_t$ such that there is an isomorphism $\pi_j : T \to S_j$ for some $j > n$. Actually, the tree and isomorphisms form a completely iterative algebra [Mo08].

Summing up, **ZFA** is proof-theoretically strong, so **ZFA** can distinguish the well-founded (WF) and non-WF parts of the universe. The set of natural numbers ω, which is a member of the WF part, is constructed by induction only, and co-inductive objects are in another partition, that is, the non-WF part. Therefore co-inductive construction does not give any effect to ω. In truth theories, the model domain only consists of natural numbers, which are constructed inductively. Co-inductive construction, which is possible by the fixed point lemma and the truth predicate, is not possible without induction, and their mixture seems to involve the existence of non-standard natural numbers.

Bibliography

[BE87] Jon Barwise. John Etchemendy. The Liar: An Essay in Truth and Circularity. Oxford University Press (1987)

[BM96] Jon Barwise, Lawrence Moss. Vicious Circles. CSLI publications (1996)

[C93] Thierry Coquand. Infinite Objects in Type Theory. TYPES (1993) 62-78.

[HH05] Volker Halbach and Leon Horsten. The deflationist's axioms for truth. in J.C. Beal and B. Armour-Garb (eds.) Deflationism and Paradox. Oxford Universoty Press.

[L01] Hannes Leitgeb. Theories of truth which have no standard models. Studia Logica, 68 (2001) 69-87.

[Mc85] Vann McGee. How truthlike can a predicate be? A negative result. Journal of Philosophical Logic, 17 (1985): 399-410.

[Mo08] Lawrence S. Moss. Coalgebra and Circularity. Preprint (2008)

[P97] Graham Priest. Yablo's paradox. Analysis 57 (1997) pp.236-242.

[V04] Matteo Viale. The cumulative hierarchy and the constructible universe of ZFA. Mathematical Logic Quarterly, 2004, volume 50, no. 1, pp.99-103.

[Yab93] Stephen Yablo. Paradox Without Self-Reference. Analysis 53 (1993) 2251-52.

[Yab06] Stephen Yablo. Circularity and Paradox. Self-reference. in Bolander, Hendricks, Pedersen (eds.), Self-reference, Stanford: CSLI Publications.

[Yat12a] Shunsuke Yatabe. A constructive naive set theory and infinity. Preprint.

[Yat12b] Shunsuke Yatabe. Yablo's paradox, a coinductive language and its semantics. Preprint.

Husserl's philosophy of arithmetic, a discussion

Roman Kossak
City University of New York

Dallas Willard in the introduction to his translation of Edmund Husserl's *Philosophy of Arithmetic* writes: "There are three main questions which Husserl addresses in his earliest writings:

1. What is number itself?
2. In what kind of cognitive act is number itself actually present in our minds?
3. How do symbols and symbolic systems used in arithmetical thought enable us to represent, and to arrive at knowledge of, number and number relations that are not "... intuitively given to our minds ...".

These questions are fundamental, but Husserl himself warns us: "The 'philosophy of arithmetic' ... does not claim to construct a thoroughgoing system of this boundary discipline, of equal importance to the mathematician and to the philosopher. Rather, in a sequence of 'psychological and logical investigations,' it claims to prepare the scientific foundations for a future construction of that discipline. In the present state of the science, nothing more than such a 'preparation' could be attempted. I would not know how to indicate even *one* question of consequence where the response could sustain a merely passable harmony among the investigators concerned."

Frege was very critical of the book, and Husserl himself, preoccupied with more general tasks of phenomenology, never returned to his early work. Nevertheless, *Philosophy of Arithmetic* is an important contribution, and it has been recently made more accessible by Stefania Cantrone's analysis in *Logic and Philosophy of Mathematics in the Early Husserl*, Springer, Synthese Library, 345, 2010.

As an introduction to the discussion, I will present excerpts from *Philosophy of Arithmetic* and other texts.