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) $P(\ulcorner A \urcorner) \rightarrow A$

(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) $\vdash_{\mathsf{Z}} P(\ulcorner D \urcorner) \rightarrow D$	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 sytnax, 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 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.

[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.

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 $\mathcal{C}$.

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) $az \equiv \top$

---

[2] In both form and motivation, $\pi uv \equiv \top$ can thus be compared to statements of the form $\mathrm{Proof}(\overline{n}, \ulcorner\varphi(x)\urcorner)$ where $\mathrm{Proof}(x, y)$ 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 $\lambda xyz.xzy(\lambda w.\top)z$.

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.