Self-Reference in Arithmetic

Volker Halbach

12th February 2021

‘Self-Reference in Arithmetic I’ (with Albert Visser), *Review of Symbolic Logic* 7 (2014), 671–691

‘Self-Reference in Arithmetic II’ (with Albert Visser), *Review of Symbolic Logic* 7 (2014), 692–712

… and in one file: *Self-Reference in Arithmetic*, `http://www.phil.uu.nl/preprints/lgps/number/316`

‘The Henkin sentence’ (with Albert Visser), in *The Life and Work of Leon Henkin* (*Essays on His Contributions*), María Manzano, Ildiko Sain and Enrique Alonso (eds.), Studies in Universal Logic, Birkhäuser, Basel, 2014, 249–264

Intensional version of the first incompleteness theorem

The sentence that states its own unprovability isn't provable.

> *Wir haben also einen Satz vor uns, der seine eigene Unbeweisbarkeit behauptet.* Gödel (1931, p. 175)
>
> *We thus have a sentence before us that claims its own unprovability.*

Löb's solution of Henkin's problem

The sentence that states its own provability is provable.

Σ_n-truth teller

For any $n \geq 1$, the Σ_n-sentence stating its own Σ_n-truth is refutable in arithmetic.

Π_n-truth teller

The Π_n-sentence stating its own Π_n-truth is provable in arithmetic ($n \geq 1$).

Löb's solution of Henkin's problem

The sentence that states its own provability is provable.

Σ_n-truth teller

For any $n \geq 1$, the Σ_n-sentence stating its own Σ_n-truth is refutable in arithmetic.

Π_n-truth teller

The Π_n-sentence stating its own Π_n-truth is provable in arithmetic ($n \geq 1$).

Löb's solution of Henkin's problem

The sentence that states its own provability is provable.

Σ_n-truth teller

For any $n \geq 1$, the Σ_n-sentence stating its own Σ_n-truth is refutable in arithmetic.

Π_n-truth teller

The Π_n-sentence stating its own Π_n-truth is provable in arithmetic ($n \geq 1$).

Löb's sentence

The sentence that says of itself that its provability implies φ is ??.

Curry

The sentence that says of itself that its truth implies φ is ??.

Visser–Yablo variants

The sentence that says of itself that all following sentences it are not true is ??.

Logical validity non-paradox

The sentence that says of itself that it is logically invalid is arithmetically provable and not logically valid.

Löb's sentence

The sentence that says of itself that its provability implies φ is ??.

Curry

The sentence that says of itself that its truth implies φ is ??.

Visser–Yablo variants

The sentence that says of itself that all following sentences it are not true is ??.

Logical validity non-paradox

The sentence that says of itself that it is logically invalid is arithmetically provable and not logically valid.

Löb's sentence

The sentence that says of itself that its provability implies φ is ??.

Curry

The sentence that says of itself that its truth implies φ is ??.

Visser–Yablo variants

The sentence that says of itself that all following sentences it are not true is ??.

Logical validity non-paradox

The sentence that says of itself that it is logically invalid is arithmetically provable and not logically valid.

Löb's sentence

The sentence that says of itself that its provability implies φ is ??.

Curry

The sentence that says of itself that its truth implies φ is ??.

Visser–Yablo variants

The sentence that says of itself that all following sentences it are not true is ??.

Logical validity non-paradox

The sentence that says of itself that it is logically invalid is arithmetically provable and not logically valid.

The analysandum: self-predication

Let P be a property such as *unprovability*, and p a corresponding adjective such as *unprovable*.

- claims its own P
- ascribes property P to itself
- predicates P of itself
- states its own P
- says of itself that it is p
- claims that it is p

An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself, and a detailed account of the library in which he is lost [...] He still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say, 'This place is aisle five, floor six, of Main Library, Stanford. I am Rudolf Lingens.'

Perry (1977, p. 492)

1. Lingens says that the man in aisle five, floor six, of Main Library, Stanford is in Stanford.
2. Lingens says about the man in aisle five, floor six, of Main Library, Stanford that he is in Stanford.
3. Lingens says about himself that he is in Stanford.
4. Lingens ascribes to himself the property of being in Stanford.

1. Lingens says that the man in aisle five, floor six, of Main Library, Stanford is in Stanford.
2. Lingens says about the man in aisle five, floor six, of Main Library, Stanford that he is in Stanford.
3. Lingens says about himself that he is in Stanford.
4. Lingens ascribes to himself the property of being in Stanford.

1. Lingens says that the man in aisle five, floor six, of Main Library, Stanford is in Stanford.
2. Lingens says about the man in aisle five, floor six, of Main Library, Stanford that he is in Stanford.
3. Lingens says about himself that he is in Stanford.
4. Lingens ascribes to himself the property of being in Stanford.

There are three steps in the construction of a self-referential formulae, that is, a formula saying about itself that it has property P:

1. Fix a Gödel coding.
2. Pick a formula expressing the property P (under the chosen coding).
3. Construct a formula ascribing to itself property P via the chosen formula.

Corresponding to these three steps there are three dimensions of intensionality. Results may be sensitive to

1. the chosen coding.
2. the formulae used to express properties (under the chosen coding).
3. the construction of a self-referential sentence from this formula.

The language, the theory etc. are kept fixed.

To show that the sentences ascribing to themselves property P are provable (refutable, true etc.) we prove the result for a specific formal sentence and then show that the result is invariant under

1. all reasonable codings,
2. all reasonable choices of formulae for expressing property P,
3. all reasonable choices of the constructions for self-reference.

This may justify the singular 'the sentence'.

For this class I concentrate on 3 and its interaction with 2.

I look at provability first and then at truth.

Henkin (1952) asked:

> *If Σ is any standard formal system adequate for recursive number theory, a formula (having a certain integer q as its Gödel number) can be constructed which expresses the proposition that the formula with Gödel number q is provable in Σ. Is this formula provable or independent in Σ?*

Kreisel (1953) replied:

> *We shall show below that the answer to Henkin's question depends on which formula is used to 'express' the notion of provability in Σ.*

Henkin (1952) asked:

> *If Σ is any standard formal system adequate for recursive number theory, a formula (having a certain integer q as its Gödel number) can be constructed which expresses the proposition that the formula with Gödel number q is provable in Σ. Is this formula provable or independent in Σ?*

Kreisel (1953) replied:

> *We shall show below that the answer to Henkin's question depends on which formula is used to 'express' the notion of provability in Σ.*

Kreisel's criterion for the expression of provability

A formula $\mathrm{Bew}(x)$ is said to express provability in Σ if it satisfies the following condition: for numerals $\overline{n}$, $\mathrm{Bew}(\overline{n})$ can be proved in Σ if and only if the formula with number n can be proved in Σ.[1]

This means that a formula is a provability predicate iff it weakly represents provability.

[1] This is the third paragraph of Kreisel's 1953 paper with the notation adapted.

A Henkin sentence is a sentence γ that *says of itself* that it's provable.

A fixed point of a formula $\varphi(x)$ (relative to a system Σ) is a formula γ such that $\Sigma \vdash \gamma \leftrightarrow \varphi(\ulcorner\gamma\urcorner)$ obtains.

To be a Henkin sentence, a sentence γ has to be at least a fixed point of the provability predicate. So if γ is a Henkin sentence we have:

$$\Sigma \vdash \gamma \leftrightarrow \mathrm{Bew}(\ulcorner\gamma\urcorner)$$

A Henkin sentence is a sentence γ that *says of itself* that it's provable.

A fixed point of a formula $\varphi(x)$ (relative to a system Σ) is a formula γ such that $\Sigma \vdash \gamma \leftrightarrow \varphi(\ulcorner\gamma\urcorner)$ obtains.

To be a Henkin sentence, a sentence γ has to be at least a fixed point of the provability predicate. So if γ is a Henkin sentence we have:

$$\Sigma \vdash \gamma \leftrightarrow \mathrm{Bew}(\ulcorner\gamma\urcorner)$$

Observation

For any given formula $\varphi(x)$ there is no formula $\chi(x)$ that defines the set of fixed points of $\varphi(x)$, that is, there is no $\chi(x)$ satisfying the following condition:

$$\mathbb{N} \vDash \chi(\ulcorner\psi\urcorner) \leftrightarrow \big(\varphi(\ulcorner\psi\urcorner) \leftrightarrow \psi\big)$$

Moreover, for any given $\varphi(x)$ the set of its provable fixed points, that is, the set of all sentences ψ with

$$\Sigma \vdash \varphi(\ulcorner\psi\urcorner) \leftrightarrow \psi$$

is not recursive but only recursively enumerable.

I suspect that Kreisel and Henkin implicitly agreed on a criterion for self-reference along the following lines, because they realized that there are trivial fixed points such as $0 = 0$:

Kreisel–Henkin criterion for self-reference

Let a formula $\varphi(x)$ expressing a certain property P in Σ be given. Then a formula γ *says about itself that it has property P* iff it is of the form $\varphi(t)$ for some closed term t that has (the code of) $\varphi(t)$ as its value.

If the usual Gödel sentence is constructed in a language with suitable function symbols, it will satisfy this condition.

In Albert's coding with built-in self-reference the Kreisel–Henkin criterion can be satisfied for arbitrary formulae without proving the usual diagonal lemma in the language.

For each formula $\varphi(x)$ there is a number n such that n is the code of $\varphi(\overline{n})$.

Kreisel's observation

There is a formula $\mathrm{Bew}_1(x)$ and a term t_1 such that the following three conditions are satisfied:

(i) Bew_1 weakly represents provability in Σ.

(ii) $\Sigma \vdash t_1 = \ulcorner \mathrm{Bew}_1(t_1) \urcorner$

(iii) $\Sigma \vdash \mathrm{Bew}_1(t_1)$

Similarly, there is a provability predicate $\mathrm{Bew}_2(x)$ and a term t_2 such that

(i) Bew_2 weakly represents provability in Σ.

(ii) $\Sigma \vdash t_2 = \ulcorner \mathrm{Bew}_2(t_2) \urcorner$

(iii) $\Sigma \vdash \neg \mathrm{Bew}_2(t_2)$

Example	
(A)	(A) = (A).

Example	
(B)	(B) = (A).

(A) is true; (B) is false.

Do (A) and (B) both say about themselves that they are identical with (A)?

Do (A) and (B) both ascribe to themselves the property of being identical with (A)?

Proof (Kreisel and Henkin)

Fix some predicate $\mathrm{Bew}(x)$ that weakly represents Σ-provability in Σ. By Gödel's diagonal lemma there is a term t_1 such that

(1) $$\Sigma \vdash t_1 = \ulcorner t_1 = t_1 \vee \mathrm{Bew}(t_1) \urcorner$$

Now define $\mathrm{Bew}_1(x)$ as

$$x = t_1 \vee \mathrm{Bew}(x).$$

Clearly $\Sigma \vdash t_1 = \ulcorner t_1 = t_1 \vee \mathrm{Bew}(t_1) \urcorner$ and hence (ii) holds by (1). Since

$$t_1 = t_1 \vee \mathrm{Bew}(t_1)$$

is provable in pure logic (and thus in Σ), $\mathrm{Bew}_1(t_1)$ is provable and (iii) is satisfied.

Proof

Fix some predicate $\mathrm{Bew}(x)$ that weakly represents Σ-provability in Σ. By Gödel's diagonal lemma there is a term t_2 such that

(2) $$\Sigma \vdash t_2 = \ulcorner t_2 \neq t_2 \wedge \mathrm{Bew}(t_2) \urcorner$$

Now define $\mathrm{Bew}_2(x)$ as

$$x \neq t_2 \wedge \mathrm{Bew}(x)$$

Clearly $\Sigma \vdash t_2 = \ulcorner t_2 \neq t_2 \wedge \mathrm{Bew}(t_2) \urcorner$ and hence (ii) holds by (2). Since

$$t_2 \neq t_2 \wedge \mathrm{Bew}(t_2)$$

is refutable in pure logic (and thus in Σ), $\Sigma \vdash \neg \mathrm{Bew}_2(t_2)$ and (iii) is satisfied.

Henkin and other people have complained ever since that Kreisel hadn't used the canonical provability predicate.

But nobody (except for Smoryński 1991) complained about the way Kreisel obtained the terms t_1 and t_2.

Now apply the *standard diagonal method* to $\mathrm{Bew}_2(x)$ with $\mathrm{Bew}(x)$ the canonical provability predicate to obtain a term t_3:

(i) $\Sigma \vdash t_3 = \ulcorner \mathrm{Bew}_2(t_3) \urcorner$

(ii) $\Sigma \vdash \mathrm{Bew}_2(t_3)$

Both t_2 and t_3 satisfy Kreisel's criterion for self-reference and say about themselves that they are provable in the sense of Bew_2.

Henkin and other people have complained ever since that Kreisel hadn't used the canonical provability predicate.

But nobody (except for Smoryński 1991) complained about the way Kreisel obtained the terms t_1 and t_2.

Now apply the *standard diagonal method* to $\mathrm{Bew}_2(x)$ with $\mathrm{Bew}(x)$ the canonical provability predicate to obtain a term t_3:

(i) $\Sigma \vdash t_3 = \ulcorner \mathrm{Bew}_2(t_3) \urcorner$

(ii) $\Sigma \vdash \mathrm{Bew}_2(t_3)$

Both t_2 and t_3 satisfy Kreisel's criterion for self-reference and say about themselves that they are provable in the sense of Bew_2.

Conclusion:

Whether Kreisel's 'Henkin sentence' is provable or refutable does not depend – contra Kreisel – *only* on the provability predicate; it also depends on how self-reference is obtained.

The answer to certain questions depends not only on the coding and the representing formulae, but also on how self-reference is obtained.

However, in the case of the Henkin sentences the intensionality from self-reference disappears once we consider canonical provability predicates.

Conclusion:

Whether Kreisel's 'Henkin sentence' is provable or refutable does not depend – contra Kreisel – *only* on the provability predicate; it also depends on how self-reference is obtained.

The answer to certain questions depends not only on the coding and the representing formulae, but also on how self-reference is obtained.

However, in the case of the Henkin sentences the intensionality from self-reference disappears once we consider canonical provability predicates.

Conclusion:

Whether Kreisel's 'Henkin sentence' is provable or refutable does not depend – contra Kreisel – *only* on the provability predicate; it also depends on how self-reference is obtained.

The answer to certain questions depends not only on the coding and the representing formulae, but also on how self-reference is obtained.

However, in the case of the Henkin sentences the intensionality from self-reference disappears once we consider canonical provability predicates.

Löb's theorem is the answer to Henkin's problem if the provability predicate is kept *canonical*. Assume $\mathrm{Bew}(x)$ satisfies the derivability conditions.

Lemma

Any two fixed points of $\mathrm{Bew}(v)$ are Σ-provably equivalent.

More formally:

$\Sigma \vdash \gamma_1 \leftrightarrow \mathrm{Bew}(\ulcorner\gamma_1\urcorner)$ and $\Sigma \vdash \gamma_2 \leftrightarrow \mathrm{Bew}(\ulcorner\gamma_2\urcorner)$ imply $\Sigma \vdash \gamma_1 \leftrightarrow \gamma_2$.

Löb's theorem is the answer to Henkin's problem if the provability predicate is kept *canonical*. Assume $\mathrm{Bew}(x)$ satisfies the derivability conditions.

Lemma

Any two fixed points of $\mathrm{Bew}(v)$ are Σ-provably equivalent.

More formally:
$\Sigma \vdash \gamma_1 \leftrightarrow \mathrm{Bew}(\ulcorner \gamma_1 \urcorner)$ and $\Sigma \vdash \gamma_2 \leftrightarrow \mathrm{Bew}(\ulcorner \gamma_2 \urcorner)$ imply $\Sigma \vdash \gamma_1 \leftrightarrow \gamma_2$.

But can one obtain a refutable Henkin sentence with canonical diagonalisation but a nonstandard provability predicate?

Theorem (Visser)

There is a provability predicate $\mathrm{Bew}^V(x)$ weakly representing provability in Σ such that its fixed point obtained by the usual diagonal construction is refutable.

Observation (Picollo)

There is a provability predicate $\mathrm{Bew}^P(x)$ weakly representing provability in Σ such that its fixed point obtained by the usual diagonal construction is neither provable nor refutable.

But can one obtain a refutable Henkin sentence with canonical diagonalisation but a nonstandard provability predicate?

Theorem (Visser)

There is a provability predicate $\mathrm{Bew}^V(x)$ weakly representing provability in Σ such that its fixed point obtained by the usual diagonal construction is refutable.

Observation (Picollo)

There is a provability predicate $\mathrm{Bew}^P(x)$ weakly representing provability in Σ such that its fixed point obtained by the usual diagonal construction is neither provable nor refutable.

Let d be the canonical fixed point operator that maps any formula $\varphi(x)$ to its Gödel fixed point and $\underset{\cdot}{d}$ its representation in Σ.

Let $\mathrm{Bew}(x)$ be some formula representing provability and construct a formula $\mathrm{Bew}^V(x)$ using some fixed point construction:

(3) $$\Sigma \vdash \mathrm{Bew}^V(x) \leftrightarrow x \neq \underset{\cdot}{d}(\ulcorner \mathrm{Bew}^V(x) \urcorner) \wedge \mathrm{Bew}(x)$$

Now apply the canonical d to the predicate $\mathrm{Bew}^V(x)$.

(i) $\Sigma \vdash \neg d(\mathrm{Bew}^V)$

(ii) $\mathrm{Bew}^V(x)$ weakly represents provability.

Henkin sentences: summary

- If a canonical provability predicate (at least one satisfying the Löb conditions) is chosen, all fixed points of this predicate are equivalent.
- There are provability predicates that have refutable and provable Henkin sentences (that are self-referential in the sense of the Kreisel–Henkin criterion).
- There is a refutable Henkin sentence obtained via the *canonical* Gödel diagonalisation method.

Once a reasonable provability predicate (along with a reasonable coding) is fixed, the intensionality from self-reference disappears, because all fixed point of provability behave in the same way.

But canonical provability is very special. For other predicates fixed points shouldn't be expected to be equivalent. Thus results about self-referential sentences are 'more intensional' for other predicates.

Other properties and formulae expressing them behave less extensionally. This is obvious for a sentence that says about itself that it's Gödel number is even.

$\sigma(x)$ is a truth predicate for Σ_n iff for all sentences $\varphi \in \Sigma_n$:

$$\Sigma \vdash \sigma(\ulcorner\varphi\urcorner) \leftrightarrow \varphi$$

In addition we may require that the compositional axioms hold for Σ_n.

For each $n \geq 1$ there is a Σ_n-truth predicate, which is Σ_n.

$0 = 0$ and $0 \neq 0$ are fixed points of each of these partial truth predicates.

We can play the same tricks as in the case of provability:

Theorem

There is a truth predicate $\sigma_n(x)$ for the set of Σ_n-sentences so that the truth teller formulated with $\sigma_n(x)$ using the standard diagonal function d is refutable in PA and provable with another construction for self-reference (with the Kreisel–Henkin property).

Theorem (Visser)

There is a truth predicate $\sigma_n(x)$ for the set of Σ_n-sentences so that the truth teller formulated with $\sigma_n(x)$ using the standard diagonal function d is provable in PA. There is also a truth predicate $\sigma'_n(x)$ for the set of Σ_n-sentences so that the truth teller formulated with $\sigma'_n(x)$ and standard diagonalisation d is refutable in PA.

Theorem

For each $n \geq 1$ the 'natural' Σ_n-truth teller is refutable and the 'natural' Π_n-truth teller is refutable in PA, if the coding is monotone.

Observation (McGee)

$\mathrm{Bew}_{\restriction\Sigma_1}$ is a Σ_1-truth predicate in PA.

The Henkin sentence obtained from $\mathrm{Bew}_{\restriction\Sigma_1}$ using the standard diagonal function is provable by Löb's theorem.

Theorem

For each $n \geq 1$ the 'natural' Σ_n-truth teller is refutable and the 'natural' Π_n-truth teller is refutable in PA, if the coding is monotone.

Observation (McGee)

$\mathrm{Bew}_{\restriction\Sigma_1}$ is a Σ_1-truth predicate in PA.

The Henkin sentence obtained from $\mathrm{Bew}_{\restriction\Sigma_1}$ using the standard diagonal function is provable by Löb's theorem.

Whether Σ_1-truth teller is refutable or not depends on the coding, the formulae expressing Σ_1-truth and the construction for self-reference.

There are also examples in axiomatic theories of truth, e.g., the consistency of $T\ulcorner\neg Tt\urcorner \leftrightarrow \neg Tt$ for all closed terms t.

The discussion about Yablo's paradox and self-reference requires an account of self-reference.

Whether Σ_1-truth teller is refutable or not depends on the coding, the formulae expressing Σ_1-truth and the construction for self-reference.

There are also examples in axiomatic theories of truth, e.g., the consistency of $T\ulcorner\neg Tt\urcorner \leftrightarrow \neg Tt$ for all closed terms t.

The discussion about Yablo's paradox and self-reference requires an account of self-reference.

Whether Σ_1-truth teller is refutable or not depends on the coding, the formulae expressing Σ_1-truth and the construction for self-reference.

There are also examples in axiomatic theories of truth, e.g., the consistency of $T\ulcorner\neg Tt\urcorner \leftrightarrow \neg Tt$ for all closed terms t.

The discussion about Yablo's paradox and self-reference requires an account of self-reference.

In the case of Henkin's problem and the intensional version of Gödel's first incompleteness theorem there is reasonable hope that we can prove invariance results.

Other cases such as truth tellers look more challenging. Saying what a reasonable coding, a reasonable representation of the property and a reasonable construction for self-reference is much harder in other cases.

Perhaps we should remain sceptical about the possibility of representing full self-reference in arithmetic; and the claims at the beginning are of only a heuristic value at best.

In the case of Henkin's problem and the intensional version of Gödel's first incompleteness theorem there is reasonable hope that we can prove invariance results.

Other cases such as truth tellers look more challenging. Saying what a reasonable coding, a reasonable representation of the property and a reasonable construction for self-reference is much harder in other cases.

Perhaps we should remain sceptical about the possibility of representing full self-reference in arithmetic; and the claims at the beginning are of only a heuristic value at best.

Can all the results (like refutable Henkin sentences) always be obtained by tweaking the representing formula while canonical diagonalisation is retained?

Is there two fixed points of a 'natural' predicate that both satisfy the Kreisel–Henkin criterion but differ in their properties?

Are there any reasonable additional or alternative conditions on top and above the Kreisel–Henkin condition?

Can all the results (like refutable Henkin sentences) always be obtained by tweaking the representing formula while canonical diagonalisation is retained?

Is there two fixed points of a ‘natural’ predicate that both satisfy the Kreisel–Henkin criterion but differ in their properties?

Are there any reasonable additional or alternative conditions on top and above the Kreisel–Henkin condition?

Definition

A fixed-point operator is a function f from the set of formulae with the variable v free into the set of formulae such that $\Sigma \vdash f(\varphi) \leftrightarrow \varphi(\ulcorner f(\varphi) \urcorner)$.

Definition

A fixed-point operator is *Kreiselian* iff for each φ with v free $f(\varphi)$ is of the form $\varphi(t)$ for some term t with $\Sigma \vdash t = \ulcorner f(\varphi) \urcorner$ (i.e. $\Sigma \vdash t = \ulcorner \varphi(t) \urcorner$).

Definition

A fixed-point operator d is *uniform* iff for each φ with v free: $d(\varphi)$ is of the form $\varphi(\underset{\cdot}{d}\ulcorner \varphi \urcorner)$, where $\underset{\cdot}{d}$ represents the function d.

Cf. Heck's (2007, p. 9) Structural Diagonal Lemma.

Kreisel's diagonalization method for obtaining a refutable Henkin sentence is not uniform.

Observation

Let the coding be monotone, t be some term and $\varphi(v, v)$ a formula with two marked free occurrences of the variable v. If d is a diagonal operator with $d(\varphi(t, x)) = \varphi(t, t)$, then d is not uniform.

Proof. Assume d is uniform: $d(\varphi(t, x)) \;=\; \varphi(t, \underset{\cdot}{d}\ulcorner\varphi(t, x)\urcorner)$. From the assumption $d(\varphi(t, x)) = \varphi(t, t)$ we obtain

$$t \;=\; \underset{\cdot}{d}\ulcorner\varphi(t, x)\urcorner$$

which contradicts monotonicity. □

What should we try to prove if we ask about the sentence that says about itself that it is p?

References

Solomon Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, 49:35–91, 1960.

Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik*, 38:173–198, 1931.

Richard Heck. Self-reference and the Languages of Arithmetic. *Philosophia Mathematica*, 15: 1–29, 2007.

Leon Henkin. A problem concerning provability. *Journal of Symbolic Logic*, 17:160, 1952.

Leon Henkin. Review of G. Kreisel: On a problem of Henkin's. *Journal of Symbolic Logic*, 19: 219–220, 1954.

Georg Kreisel. On a problem of Henkin's. *Indagationes Mathematicae*, 15:405–406, 1953.

Peter Milne. On Gödel Sentences and What They Say. *Philosophia Mathematica*, 15:193–226, 2007.

John Perry. Frege on demonstratives. *Philosophical Review*, 86:474–497, 1977.

Craig Smoryński. The development of self-reference: Löb's theorem. In Thomas Drucker, editor, *Perspectives on the History of Mathematical Logic*, pages 110–133. Birkhäuser, Boston, 1991.