TRUTH TELLERS

Volker Halbach

Scandinavian Logic Symposium
Tampere

25th August 2014

I’m wrote two papers with Albert Visser on this and related topics:

Self-Reference in Arithmetic,
`http://www.phil.uu.nl/preprints/lgps/number/316`
to appear as *Self-reference in Arithmetic* I and *Self-reference in Arithmetic* II in the *Review of Symbolic Logic*

The Henkin sentence, The Life and Work of Leon Henkin (Essays on His Contributions), María Manzano, Ildiko Sain and Enrique Alonso (eds), Studies in Universal Logic, Birkhäuser, to appear

Albert doesn’t agree with all my philosophical claims here.

A *truth teller sentence* is a sentence that says of itself that it's true.

I'm interested in truth tellers in formal languages, in particular, the language of arithmetic possibly augmented with a new predicate symbol for truth.

I assume that we have function symbols at least for certain primitive recursive functions in the language, in particular those expressing substitution, taking the numeral of a number etc.

I write $\ulcorner\phi\urcorner$ for the numeral of the code of the expression ϕ. Unless otherwise stated, the coding is not fancy.

A *truth teller sentence* is a sentence that says of itself that it's true.

I'm interested in truth tellers in formal languages, in particular, the language of arithmetic possibly augmented with a new predicate symbol for truth.

I assume that we have function symbols at least for certain primitive recursive functions in the language, in particular those expressing substitution, taking the numeral of a number etc.

I write $\ulcorner\phi\urcorner$ for the numeral of the code of the expression ϕ. Unless otherwise stated, the coding is not fancy.

I consider the following notions of truth and approximations to truth:

- truth as a primitive notion
- partial truth predicates: Tr_{Σ_n}, Tr_{Π_n}, $\mathrm{Bew}_{\restriction\Sigma_1}$ in PA

Assume that a formula $\tau(x)$ is fixed as truth predicate. Which sentences do say about themselves that they are true (in the sense of $\tau(x)$)?

If γ says about itself that it is true then γ will be a *fixed point* of $\tau(x)$, that is,

- $\Sigma \vdash \gamma \leftrightarrow \tau(\ulcorner \gamma \urcorner)$, where Σ is your favourite system, or at least
- $\mathbb{N} \vDash \gamma \leftrightarrow \tau(\ulcorner \gamma \urcorner)$

But being a fixed point isn't sufficient for being a truth teller.

Example: $\Sigma \vdash 0 = 0 \leftrightarrow \tau(\ulcorner 0 = 0 \urcorner)$ or $\Sigma \vdash 0 \neq 0 \leftrightarrow \tau(\ulcorner 0 \neq 0 \urcorner)$

Assume that a formula $\tau(x)$ is fixed as truth predicate. Which sentences do say about themselves that they are true (in the sense of $\tau(x)$)?

If γ says about itself that it is true then γ will be a *fixed point* of $\tau(x)$, that is,

- $\Sigma \vdash \gamma \leftrightarrow \tau(\ulcorner\gamma\urcorner)$, where Σ is your favourite system, or at least
- $\mathbb{N} \vDash \gamma \leftrightarrow \tau(\ulcorner\gamma\urcorner)$

But being a fixed point isn't sufficient for being a truth teller.

Example: $\Sigma \vdash 0=0 \leftrightarrow \tau(\ulcorner 0=0\urcorner)$ or $\Sigma \vdash 0\neq 0 \leftrightarrow \tau(\ulcorner 0\neq 0\urcorner)$

OBSERVATION

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

$$\mathbb{N} \models \chi(\ulcorner\psi\urcorner) \leftrightarrow \big(\tau(\ulcorner\psi\urcorner) \leftrightarrow \psi\big)$$

Moreover, for any given $\tau(x)$ the set of its Σ-provable fixed points (Σ must prove diagon.), that is, the set of all sentences ψ with

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

is not recursive but only recursively enumerable.

Only in very special cases will all fixed points be equivalent.

Let $sub(y, z)$ be a function expression representing naturally the function that substitutes the numeral of z for the fixed variable x in y.

Let g be term $sub(\ulcorner\tau(sub(x,x))\urcorner, \ulcorner\tau(sub(x,x))\urcorner)$

$I\Sigma_1 \vdash g = \ulcorner\tau(sub(\ulcorner\tau(sub(x,x))\urcorner, \ulcorner\tau(sub(x,x))\urcorner))\urcorner$

$\tau(g)$ is a truth teller, the *canonical* truth teller.

Let $sub(y, z)$ be a function expression representing naturally the function that substitutes the numeral of z for the fixed variable x in y.

Let g be term $sub(\ulcorner\tau(sub(x, x))\urcorner, \ulcorner\tau(sub(x, x))\urcorner)$

$I\Sigma_1 \vdash g = \ulcorner\tau(sub(\ulcorner\tau(sub(x, x))\urcorner, \ulcorner\tau(sub(x, x))\urcorner))\urcorner$

$\tau(g)$ is a truth teller, the *canonical* truth teller.

Let $sub(y, z)$ be a function expression representing naturally the function that substitutes the numeral of z for the fixed variable x in y.

Let g be term $sub(\ulcorner\tau(sub(x, x))\urcorner, \ulcorner\tau(sub(x, x))\urcorner)$

$I\Sigma_1 \vdash g = \ulcorner\tau(sub(\ulcorner\tau(sub(x, x))\urcorner, \ulcorner\tau(sub(x, x))\urcorner))\urcorner$

$\tau(g)$ is a truth teller, the *canonical* truth teller.

DEFINITION

Assume again that a truth predicate $\tau(x)$ is fixed.
Then γ is a *KH-truth teller* iff γ is of the form $\tau(t)$ and $\mathsf{I}\Sigma_1 \vdash t = \ulcorner \tau(t) \urcorner$.

OBSERVATION

If $\tau(t)$ is a KH-truth teller, then, obviously,
$\mathsf{I}\Sigma_1 \vdash \tau(t) \leftrightarrow \tau(\ulcorner \tau(t) \urcorner)$, that is, $\tau(t)$ is a $\mathsf{I}\Sigma_1$-provable fixed point of τ.

'KH' stands for 'Kreisel–Henkin'. Cf. Henkin (1952); Kreisel (1953); Henkin (1954).

DEFINITION

Assume again that a truth predicate $\tau(x)$ is fixed.
Then γ is a *KH-truth teller* iff γ is of the form $\tau(t)$ and $\mathrm{I}\Sigma_1 \vdash t = \ulcorner\tau(t)\urcorner$.

OBSERVATION

If $\tau(t)$ is a KH-truth teller, then, obviously, $\mathrm{I}\Sigma_1 \vdash \tau(t) \leftrightarrow \tau(\ulcorner\tau(t)\urcorner)$, that is, $\tau(t)$ is a $\mathrm{I}\Sigma_1$-provable fixed point of τ.

'KH' stands for 'Kreisel–Henkin'. Cf. Henkin (1952); Kreisel (1953); Henkin (1954).

Add a new unary predicate symbol T to the language of arithmetic.

Our $\tau(x)$ is now the formula Tx.

There are many ways to obtain an interpretation or axiomatization for this language, such that T is characterized as a truth predicate (in some sense).

I look at a special case of the semantics in Kripke (1975).

A set S of sentences is an SK-Kripke set iff S doesn't contain any sentence together with its negation and is closed under the following conditions, where s and t are closed terms:

- $\text{value}(s) = \text{value}(t) \;\Rightarrow\; (s = t) \in S$
- $\text{value}(s) \neq \text{value}(t) \;\Rightarrow\; (\neg s = t) \in S$
- $\phi \in S \;\Rightarrow\; (\neg\neg\phi) \in S$
- $\phi, \psi \in S \;\Rightarrow\; (\phi \wedge \psi) \in S$
- $\neg\phi \in S$ or $\neg\psi \in S \;\Rightarrow\; (\neg(\phi \wedge \psi)) \in S$
- $\phi(t) \in S$ for all closed terms $t \;\Rightarrow\; (\forall v \phi(v)) \in S$ (renam. var.)
- $(\neg\phi(t)) \in S$ for some closed term $t \;\Rightarrow\; (\neg\forall v \phi(v)) \in S$
- $\phi \in S$ and $\text{value}(t) = \ulcorner\phi\urcorner \;\Rightarrow\; (Tt) \in S$
- $(\neg\phi) \in S$ and $\text{value}(t) = \ulcorner\phi\urcorner \;\Rightarrow\; (\neg Tt) \in S$

OBSERVATION

There are SK-Kripke sets that contain the canonical truth teller, other SK-Kripke sets that contain its negation, still other SK-Kripke sets that contain neither. The same SK-Kripke set can contain a KH-truth teller and the negation of another KH-truth teller.

In most axiomatic truth theories no truth teller is decided (exception KFB).

Example PUTB with the characteristic axiom schema

$$\forall t\,\big(T\phi(t) \leftrightarrow \phi(\mathrm{value}(t))\big)$$

where $\phi(x)$ is positive in T.

OBSERVATION

There are SK-Kripke sets that contain the canonical truth teller, other SK-Kripke sets that contain its negation, still other SK-Kripke sets that contain neither. The same SK-Kripke set can contain a KH-truth teller and the negation of another KH-truth teller.

In most axiomatic truth theories no truth teller is decided (exception KFB).

Example PUTB with the characteristic axiom schema

$$\forall t \left(T\phi(t) \leftrightarrow \phi(\text{value}(t)) \right)$$

where $\phi(x)$ is positive in T.

Conclusion If truth is treated as a primitive notions and one postulates just basic disquotational features for this notion, truth tellers cannot be decided.

A formula is Σ_0 (and also Π_0) iff it doesn't contain any unbounded quantifiers. A formula is Σ_{n+1} iff it is of the form $\exists\vec{x}\,\phi$, where ϕ is Π_n or obtained from such formula by combining them using conjunction and disjunction.

OBSERVATION

For each $n > 0$ there is a Σ_n-formula σ_n such that the following holds for all Σ_n-sentences ψ:

$$\mathsf{PA} \vdash \sigma_n(\ulcorner\psi\urcorner) \leftrightarrow \psi$$

Such formulae σ_n are called Σ_n-truth predicates. (Note that they may not have higher complexity than Σ_n).

For Π_n an analogous claim holds.

A formula is Σ_0 (and also Π_0) iff it doesn't contain any unbounded quantifiers. A formula is Σ_{n+1} iff it is of the form $\exists\vec{x}\,\phi$, where ϕ is Π_n or obtained from such formula by combining them using conjunction and disjunction.

OBSERVATION

For each $n > 0$ there is a Σ_n-formula σ_n such that the following holds for all Σ_n-sentences ψ:

$$\mathsf{PA} \vdash \sigma_n(\ulcorner\psi\urcorner) \leftrightarrow \psi$$

Such formulae σ_n are called Σ_n-truth predicates. (Note that they may not have higher complexity than Σ_n).

For Π_n an analogous claim holds.

OBSERVATION

Assume that $n > 0$, σ_n is a Σ_n-truth predicate and τ is a KH-truth teller. Then τ is Σ_n. I call such τ Σ_n-*truth tellers*. An analogous claim holds for Π_n-truth tellers.

Proof If τ is a KH-truth teller, then it is of the form $\sigma_n(t)$ with $\mathsf{PRA} \vdash t = \ulcorner \sigma_n(t) \urcorner$. Clearly, $\sigma_n(t)$ is Σ_n.

OBSERVATION

For each n there are provable, refutable and independent Π_n- and Σ_n-truth tellers.

The behaviour of a Σ_n-truth teller depends on at least three factors:

- the chosen coding
- the Σ_n-truth predicate
- the way the KH-truth teller is obtained

I'll now demonstrate the effects of tweaking any of these factors.

OBSERVATION

For each n there are provable, refutable and independent Π_n- and Σ_n-truth tellers.

The behaviour of a Σ_n-truth teller depends on at least three factors:

- the chosen coding
- the Σ_n-truth predicate
- the way the KH-truth teller is obtained

I'll now demonstrate the effects of tweaking any of these factors.

OBSERVATION

For each n there are provable, refutable and independent Π_n- and Σ_n-truth tellers.

The behaviour of a Σ_n-truth teller depends on at least three factors:

- the chosen coding
- the Σ_n-truth predicate
- the way the KH-truth teller is obtained

I'll now demonstrate the effects of tweaking any of these factors.

OBSERVATION

For each n there are provable, refutable and independent Π_n- and Σ_n-truth tellers.

The behaviour of a Σ_n-truth teller depends on at least three factors:

- the chosen coding
- the Σ_n-truth predicate
- the way the KH-truth teller is obtained

I'll now demonstrate the effects of tweaking any of these factors.

OBSERVATION

For each n there are provable, refutable and independent Π_n- and Σ_n-truth tellers.

The behaviour of a Σ_n-truth teller depends on at least three factors:

- the chosen coding
- the Σ_n-truth predicate
- the way the KH-truth teller is obtained

I'll now demonstrate the effects of tweaking any of these factors.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

$$(\dagger) \qquad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \rightarrow \exists v < y\ \phi(v)) \qquad \text{for } \phi(v) \text{ in } \Sigma_1$$

Now reason in PA:

Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

1 + $(\dagger)$ give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

$$(\dagger) \qquad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \rightarrow \exists v < y\ \phi(v)) \qquad \text{for } \phi(v) \text{ in } \Sigma_1$$

Now reason in PA:
Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

1 + ($\dagger$) give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

(†) $\quad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \rightarrow \exists v < y\ \phi(v)) \quad$ for $\phi(v)$ in Σ_1

Now reason in PA:

Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

1 + (†) give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

$$(\dagger) \qquad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \to \exists v < y\ \phi(v)) \qquad \text{for } \phi(v) \text{ in } \Sigma_1$$

Now reason in PA:

Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

$1 + (\dagger)$ give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

$(\dagger) \quad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \rightarrow \exists v < y\ \phi(v)) \quad$ for $\phi(v)$ in Σ_1

Now reason in PA:

Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

1 + $(\dagger)$ give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

(†) $\quad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \rightarrow \exists v < y\ \phi(v)) \quad$ for $\phi(v)$ in Σ_1

Now reason in PA:

Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

1 + (†) give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

THEOREM (McGee and Visser)

Suppose we employ a standard, monotone Gödel coding and Tr_{Σ_n} the canonical Σ_n-truth predicate. If $Tr_{\Sigma_n}(t)$ is a KH-truth teller, $\mathsf{PA} \vdash \neg Tr_{\Sigma_n}(t)$ obtains for $n \geq 1$.

Proof for $n = 1$. $Tr_{\Sigma_1}(x)$ is $\exists y\, \theta(y, x)$. Thus, $\mathsf{PA} \vdash t = \ulcorner \exists y\, \theta(y, t) \urcorner$.

(†) $\quad \mathsf{PA} \vdash \forall y\, (\theta(y, \ulcorner \exists v\, \phi(v) \urcorner) \rightarrow \exists v < y\ \phi(v)) \quad$ for $\phi(v)$ in Σ_1

Now reason in PA:
Suppose $\exists y\, \theta(y, t)$. Let y_0 be the smallest witness of $\exists y\, \theta(y, t)$.

1. $\theta(y_0, t)$
2. $\forall z < y_0\ \neg\theta(z, t)$

1 + (†) give $\exists z < y_0\ \theta(z, t)$, which contradicts 2.

Similarly, canonical Π_n-truth tellers are provable.

We used the monotocity of the coding schema and the canonical definition of Tr_{Σ_1} in (†) .

OBSERVATION

$\mathsf{PA} \vdash \mathrm{Bew}_{\mathsf{I}\Sigma_1}(\ulcorner \phi \urcorner) \leftrightarrow \phi$ for all Σ_1 sentences ϕ. Thus, $\mathrm{Bew}_{\mathsf{I}\Sigma_1}$ is a Σ_1-truth predicate.

OBSERVATION

If $\mathsf{I}\Sigma_1 \vdash \mathrm{Bew}_{\mathsf{I}\Sigma_1}(\ulcorner \phi \urcorner) \leftrightarrow \phi$, then $\mathsf{I}\Sigma_1 \vdash \phi$, by Löb's theorem. Consequently all KH-truth tellers based on $\mathrm{Bew}_{\mathsf{I}\Sigma_1}$ are provable.

We used the monotocity of the coding schema and the canonical definition of Tr_{Σ_1} in $(\dagger)$.

OBSERVATION

$\mathsf{PA} \vdash \mathrm{Bew}_{\mathsf{I}\Sigma_1}(\ulcorner\phi\urcorner) \leftrightarrow \phi$ for all Σ_1 sentences ϕ. Thus, $\mathrm{Bew}_{\mathsf{I}\Sigma_1}$ is a Σ_1-truth predicate.

OBSERVATION

If $\mathsf{I}\Sigma_1 \vdash \mathrm{Bew}_{\mathsf{I}\Sigma_1}(\ulcorner\phi\urcorner) \leftrightarrow \phi$, then $\mathsf{I}\Sigma_1 \vdash \phi$, by Löb's theorem. Consequently all KH-truth tellers based on $\mathrm{Bew}_{\mathsf{I}\Sigma_1}$ are provable.

THEOREM

Let $n \geq 1$ be given and the coding be monotone. Then there is a Σ_n-truth predicate with a provable and a refutable KH-truth tellers. More explicitly, there is a Σ_n-truth predicate σ_n and terms t_1 and t_2 such that

(i) $\mathsf{PA} \vdash t_1 = \ulcorner \sigma_n(t_1) \urcorner$ and $\mathsf{PA} \vdash \sigma_n(t_1)$

(ii) $\mathsf{PA} \vdash t_2 = \ulcorner \sigma_n(t_2) \urcorner$ and $\mathsf{PA} \vdash \neg\sigma_n(t_2)$

We use a trick due to Henkin (1954) that improves a trick by Kreisel (1953).

Picollo produced also independent truth tellers.

We assume that Σ_n is closed under disjunction.

THEOREM

Let $n \geq 1$ be given and the coding be monotone. Then there is a Σ_n-truth predicate with a provable and a refutable KH-truth tellers. More explicitly, there is a Σ_n-truth predicate σ_n and terms t_1 and t_2 such that

(i) $\mathsf{PA} \vdash t_1 = \ulcorner \sigma_n(t_1) \urcorner$ and $\mathsf{PA} \vdash \sigma_n(t_1)$

(ii) $\mathsf{PA} \vdash t_2 = \ulcorner \sigma_n(t_2) \urcorner$ and $\mathsf{PA} \vdash \neg\sigma_n(t_2)$

We use a trick due to Henkin (1954) that improves a trick by Kreisel (1953).

Picollo produced also independent truth tellers.

We assume that Σ_n is closed under disjunction.

Apply the diagonal lemma to the canonical Σ_n-truth predicate $Tr_{\Sigma_n}(x)$ to obtain a term t_1 such that

(1) $$\mathsf{PA} \vdash t_1 = \ulcorner t_1 = t_1 \vee Tr_{\Sigma_n}(t_1) \urcorner$$

Now define $\sigma_n(x)$ as

(2) $$x = t_1 \vee Tr_{\Sigma_n}(x).$$

σ_n is a Σ_n-truth predicate, that is, for all ϕ in Σ_n:

(3) $$\mathsf{PA} \vdash \left(\ulcorner\phi\urcorner = t_1 \vee Tr_{\Sigma_n}(\ulcorner\phi\urcorner)\right) \leftrightarrow \phi$$

$\mathsf{PA} \vdash t_1 = t_1 \vee Tr_{\Sigma_n}(t_1)$ and (1) yield (i) of the theorem.

Applying the canonical diagonalization to σ_n yields a different term t_2 s.t.

(4) $$\mathsf{PA} \vdash t_2 = \ulcorner t_2 = t_1 \vee Tr_{\Sigma_n}(t_2) \urcorner$$

By an argument like above, we get $\mathsf{PA} \vdash \neg(t_2 = t_1 \vee Tr_{\Sigma_n}(t_2))$. This gives (ii).

Apply the diagonal lemma to the canonical Σ_n-truth predicate $Tr_{\Sigma_n}(x)$ to obtain a term t_1 such that

(1) $$\mathsf{PA} \vdash t_1 = \ulcorner t_1 = t_1 \vee Tr_{\Sigma_n}(t_1) \urcorner$$

Now define $\sigma_n(x)$ as

(2) $$x = t_1 \vee Tr_{\Sigma_n}(x).$$

σ_n is a Σ_n-truth predicate, that is, for all ϕ in Σ_n:

(3) $$\mathsf{PA} \vdash \left(\ulcorner\phi\urcorner = t_1 \vee Tr_{\Sigma_n}(\ulcorner\phi\urcorner)\right) \leftrightarrow \phi$$

$\mathsf{PA} \vdash t_1 = t_1 \vee Tr_{\Sigma_n}(t_1)$ and (1) yield (i) of the theorem.

Applying the canonical diagonalization to σ_n yields a different term t_2 s.t.

(4) $$\mathsf{PA} \vdash t_2 = \ulcorner t_2 = t_1 \vee Tr_{\Sigma_n}(t_2) \urcorner$$

By an argument like above, we get $\mathsf{PA} \vdash \neg(t_2 = t_1 \vee Tr_{\Sigma_n}(t_2))$. This gives (ii).

Apply the diagonal lemma to the canonical Σ_n-truth predicate $Tr_{\Sigma_n}(x)$ to obtain a term t_1 such that

(1) $$\mathsf{PA} \vdash t_1 = \ulcorner t_1 = t_1 \vee Tr_{\Sigma_n}(t_1) \urcorner$$

Now define $\sigma_n(x)$ as

(2) $$x = t_1 \vee Tr_{\Sigma_n}(x).$$

σ_n is a Σ_n-truth predicate, that is, for all ϕ in Σ_n:

(3) $$\mathsf{PA} \vdash \left(\ulcorner\phi\urcorner = t_1 \vee Tr_{\Sigma_n}(\ulcorner\phi\urcorner)\right) \leftrightarrow \phi$$

$\mathsf{PA} \vdash t_1 = t_1 \vee Tr_{\Sigma_n}(t_1)$ and (1) yield (i) of the theorem.

Applying the canonical diagonalization to σ_n yields a different term t_2 s.t.

(4) $$\mathsf{PA} \vdash t_2 = \ulcorner t_2 = t_1 \vee Tr_{\Sigma_n}(t_2) \urcorner$$

By an argument like above, we get $\mathsf{PA} \vdash \neg(t_2 = t_1 \vee Tr_{\Sigma_n}(t_2))$. This gives (ii).

Apply the diagonal lemma to the canonical Σ_n-truth predicate $Tr_{\Sigma_n}(x)$ to obtain a term t_1 such that

(1) $$\mathsf{PA} \vdash t_1 = \ulcorner t_1 = t_1 \vee Tr_{\Sigma_n}(t_1) \urcorner$$

Now define $\sigma_n(x)$ as

(2) $$x = t_1 \vee Tr_{\Sigma_n}(x).$$

σ_n is a Σ_n-truth predicate, that is, for all ϕ in Σ_n:

(3) $$\mathsf{PA} \vdash \left(\ulcorner\phi\urcorner = t_1 \vee Tr_{\Sigma_n}(\ulcorner\phi\urcorner)\right) \leftrightarrow \phi$$

$\mathsf{PA} \vdash t_1 = t_1 \vee Tr_{\Sigma_n}(t_1)$ and (1) yield (i) of the theorem.

Applying the canonical diagonalization to σ_n yields a different term t_2 s.t.

(4) $$\mathsf{PA} \vdash t_2 = \ulcorner t_2 = t_1 \vee Tr_{\Sigma_n}(t_2) \urcorner$$

By an argument like above, we get $\mathsf{PA} \vdash \neg(t_2 = t_1 \vee Tr_{\Sigma_n}(t_2))$. This gives (ii).

Apply the diagonal lemma to the canonical Σ_n-truth predicate $Tr_{\Sigma_n}(x)$ to obtain a term t_1 such that

(1) $$\mathsf{PA} \vdash t_1 = \ulcorner t_1 = t_1 \vee Tr_{\Sigma_n}(t_1) \urcorner$$

Now define $\sigma_n(x)$ as

(2) $$x = t_1 \vee Tr_{\Sigma_n}(x).$$

σ_n is a Σ_n-truth predicate, that is, for all ϕ in Σ_n:

(3) $$\mathsf{PA} \vdash \left(\ulcorner\phi\urcorner = t_1 \vee Tr_{\Sigma_n}(\ulcorner\phi\urcorner)\right) \leftrightarrow \phi$$

$\mathsf{PA} \vdash t_1 = t_1 \vee Tr_{\Sigma_n}(t_1)$ and (1) yield (i) of the theorem.

Applying the canonical diagonalization to σ_n yields a different term t_2 s.t.

(4) $$\mathsf{PA} \vdash t_2 = \ulcorner t_2 = t_1 \vee Tr_{\Sigma_n}(t_2) \urcorner$$

By an argument like above, we get $\mathsf{PA} \vdash \neg(t_2 = t_1 \vee Tr_{\Sigma_n}(t_2))$. This gives (ii).

Summary

1. The coding and the formula expressing truth are fixed: Whether a Σ_n-truth teller is provable or refutable (or independent) can depend on which method is used to obtain self-reference in the KH-sense.
2. The coding and the method for obtaining self-reference are fixed (in the canonical way). Whether a Σ_n-truth teller is provable or refutable (or independent) can depend on which formula is used to to express Σ_n-truth (proved here only for $n = 1$).
3. The 'way' of expressing truth and the method for obtaining self-reference are fixed. I conjecture that whether a Σ_n-truth teller is provable or refutable (or independent) can depend on the coding schema (This needs to be made precise).

Discussions on intensionality in metamathematic have focused on the kind of intensionality mentioned in 2.

Summary

1. The coding and the formula expressing truth are fixed: Whether a Σ_n-truth teller is provable or refutable (or independent) can depend on which method is used to obtain self-reference in the KH-sense.
2. The coding and the method for obtaining self-reference are fixed (in the canonical way). Whether a Σ_n-truth teller is provable or refutable (or independent) can depend on which formula is used to to express Σ_n-truth (proved here only for $n = 1$).
3. The 'way' of expressing truth and the method for obtaining self-reference are fixed. I conjecture that whether a Σ_n-truth teller is provable or refutable (or independent) can depend on the coding schema (This needs to be made precise).

Discussions on intensionality in metamathematic have focused on the kind of intensionality mentioned in 2.

Summary

1. The coding and the formula expressing truth are fixed: Whether a Σ_n-truth teller is provable or refutable (or independent) can depend on which method is used to obtain self-reference in the KH-sense.
2. The coding and the method for obtaining self-reference are fixed (in the canonical way). Whether a Σ_n-truth teller is provable or refutable (or independent) can depend on which formula is used to to express Σ_n-truth (proved here only for $n = 1$).
3. The 'way' of expressing truth and the method for obtaining self-reference are fixed. I conjecture that whether a Σ_n-truth teller is provable or refutable (or independent) can depend on the coding schema (This needs to be made precise).

Discussions on intensionality in metamathematic have focused on the kind of intensionality mentioned in 2.

A comparison with canonical provability in PA

A Henkin sentence is a sentence that say of itself that it's provable (in a fixed system; let's say PA).

If the canonical provability predicate is fixed, or any predicate satisfying the Löb derivability conditions, then all Henkin sentences behave in the same way, independently of the coding and the way the fixed points are obtained. In fact, all fixed points (whether 'self-referential' or not) behave in the same way.

- If truth is treated as a primitive predicate, one can say only very little about the truth teller.
- There are truth tellers or at least approximations to them in pure arithmetic. These sentences must be true or false.
- The behaviour of arithmetical truth tellers behaves on at least three factors: the coding, the formula expressing truth and the method used for obtaining self-reference
- In this sense the problem of truth teller is much more intensional than that of canonical provability.
- How do other properties behave? How intensional are their truth-teller? Examples: sentences saying about themselves that they are Rosser-provable.
- Given a coding schema and formula expressing truth, which sentences say about themselves that they are true? Is the KH-property sufficient?

- If truth is treated as a primitive predicate, one can say only very little about the truth teller.
- There are truth tellers or at least approximations to them in pure arithmetic. These sentences must be true or false.
- The behaviour of arithmetical truth tellers behaves on at least three factors: the coding, the formula expressing truth and the method used for obtaining self-reference
- In this sense the problem of truth teller is much more intensional than that of canonical provability.
- How do other properties behave? How intensional are their truth-teller? Examples: sentences saying about themselves that they are Rosser-provable.
- Given a coding schema and formula expressing truth, which sentences say about themselves that they are true? Is the KH-property sufficient?

- If truth is treated as a primitive predicate, one can say only very little about the truth teller.
- There are truth tellers or at least approximations to them in pure arithmetic. These sentences must be true or false.
- The behaviour of arithmetical truth tellers behaves on at least three factors: the coding, the formula expressing truth and the method used for obtaining self-reference
- In this sense the problem of truth teller is much more intensional than that of canonical provability.
- How do other properties behave? How intensional are their truth-teller? Examples: sentences saying about themselves that they are Rosser-provable.
- Given a coding schema and formula expressing truth, which sentences say about themselves that they are true? Is the KH-property sufficient?

- If truth is treated as a primitive predicate, one can say only very little about the truth teller.
- There are truth tellers or at least approximations to them in pure arithmetic. These sentences must be true or false.
- The behaviour of arithmetical truth tellers behaves on at least three factors: the coding, the formula expressing truth and the method used for obtaining self-reference
- In this sense the problem of truth teller is much more intensional than that of canonical provability.
- How do other properties behave? How intensional are their truth-teller? Examples: sentences saying about themselves that they are Rosser-provable.
- Given a coding schema and formula expressing truth, which sentences say about themselves that they are true? Is the KH-property sufficient?

- If truth is treated as a primitive predicate, one can say only very little about the truth teller.
- There are truth tellers or at least approximations to them in pure arithmetic. These sentences must be true or false.
- The behaviour of arithmetical truth tellers behaves on at least three factors: the coding, the formula expressing truth and the method used for obtaining self-reference
- In this sense the problem of truth teller is much more intensional than that of canonical provability.
- How do other properties behave? How intensional are their truth-teller? Examples: sentences saying about themselves that they are Rosser-provable.
- Given a coding schema and formula expressing truth, which sentences say about themselves that they are true? Is the KH-property sufficient?

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.

Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–712, 1975. reprinted in Martin (1984).

Robert L. Martin, editor. *Recent Essays on Truth and the Liar Paradox*. Clarendon Press and Oxford University Press, Oxford and New York, 1984.