Boundedly axiomatizable theories meet (in)completeness

Ali Enayat

Sep 2–5, 2024

WORMSHOP, GHENT

Outline of the tutorial

- Tarski's Undefinability of Truth (TUT), and the use of partial truth predicates in establishing the essential unboundedness of inductive sequential theories (such as PA and ZF).
- Rudiments of sequential theories, and the proof of incompleteness of boundedly axiomatizable seqential theories formulated in a finite language.
- More incompleteness results, and a pointer to some new completeness results.

Semantic form of TUT

- Let $\mathcal{L}$ be a language (signature), and $\mathcal{M}$ be an $\mathcal{L}$-structure. Also let $\varphi \mapsto \#(\varphi) \in M$ is an arbitrary mapping of unary $\mathcal{L}$-formulae $\varphi = \varphi(x)$ into M.

- **Theorem.** (Semantic Form of TUT) *There is no binary $\mathcal{L}$-formula $T(x, y)$ such that for all unary $\mathcal{L}$-formulae*: $\mathcal{M} \models \forall x \left(T\left(x, \#(\varphi)\right) \leftrightarrow \varphi(x)\right)$.

  Proof. Suppose not and consider $R(x) = \neg T(x, x)$. Then:

  (1) $\mathcal{M} \models \forall x \left(T\left(x, \#(R)\right) \leftrightarrow R(x)\right)$.

  If $r := \#(R)$, by (1) and the definition of R we obtain:

  (2) $\mathcal{M} \models T(r, r) \leftrightarrow R(r) \leftrightarrow \neg T(r, r),$ contradiction. $\square$

- The above proof is reminiscent of the proof of Russell's Paradox (1901), and of the proof of Cantor's theorem (1891) on nonexistence of a surjection of a set X onto $\mathcal{P}(X)$.

Kripke's general formulation

Cantor's Diagonal Principle

A relation is called *arithmetical* if it is definable in L, the language of arithmetic. Since L contains RE, it follows that all r.e. relations are arithmetical. Also, since L contains negation, it follows that all complements of r.e. relations are arithmetical. That L contains negation also implies that the enumeration theorem fails for arithmetical sets, i.e. there is no arithmetical relation that enumerates all the arithmetical relations; similarly, there is no recursive relation that enumerates all the recursive relations.

The best way to see this is by proving a general theorem. As in the enumeration theorem for r.e. sets, if R is a two-place relation, we write R_x for $\{y: R(x, y)\}$. We give the following

Definition: Let X be a set, F be a family of subsets of X, and R a two place relation defined on X. R is said to *supernumerate* F iff for any $S \in F$, there is an $x \in X$ such that $S = R_x$. R is said to *enumerate* F iff R supernumerates F and for all $x \in X$, $R_x \in F$.

The content of the enumeration theorem is thus that there is an r.e. relation which enumerates the r.e. sets. Next we have

Cantor's Diagonal Principle: The following two conditions are incompatible:

(i) R supernumerates F
(ii) The complement of the*Diagonal Set* is in F (the Diagonal Set is $\{x \in X: R(x, x)\}$).

Proof: Suppose (i)-(ii) hold. Then by (ii) $X\text{-}\{x \in X: R(x, x)\} = \{x \in X: {\sim}R(x, x)\} \in F$. By (i), $\{x \in X: {\sim}R(x, x)\} = R_y$ for some y. But then $R(y, x)$ iff ${\sim}R(x, x)$ for all $x \in X$, so in particular $R(y, y)$ iff ${\sim}R(y, y)$, contradiction.

Source: p.66 of Kripke's Lecture Notes on Elementary Recursion Theorem, Princeton, 1996

Semantic form of TUT and incompleteness

Assume $\varphi \mapsto \#(\varphi) \in \omega$, where $\varphi \in \mathsf{Sent}_{\mathcal{L}_{\mathrm{PA}}}$.

Let TA (true arithmetic) denote:

$$\{\#(\varphi) : \varphi \in \mathsf{Sent}_{\mathcal{L}_{\mathrm{PA}}}, \quad (\omega, +, \cdot) \models \varphi\}.$$

- **Corollary**. (Incompleteness of PA) TA *is not definable in* $(\omega, +, \cdot)$.
- **Corollary**. TA *is not axiomatizable by a subtheory of itself the set of whose* $\#$-codes *is definable in* $(\omega, +, \cdot)$. *Hence* PA *is incomplete*.

Semantic form of TUT, cont'd

- Let $\mathrm{Form}^n_{\mathcal{L}}$ = the set of n-ary $\mathcal{L}$-formulae, and suppose $\mathcal{M}$ is an $\mathcal{L}$-structure with a pairing function $\langle \cdot, \cdot \rangle$. Also assume that the coding $\varphi \mapsto \#(\varphi) \in M$ is 1-1. In this context, the (codes of) sentences in the elementary diagram of $\mathcal{M}$ can be split into:

$$\mathrm{ED}^{+}(\mathcal{M}) = \{\langle \#(\varphi), \langle a_1, ..., a_n \rangle \rangle \in M : \mathcal{M} \models \varphi(a_1, ..., a_n), \varphi \in \mathrm{Form}^n_{\mathcal{L}}\},$$

$$\mathrm{ED}^{-}(\mathcal{M}) = \{\langle \#(\varphi), \langle a_1, ..., a_n \rangle \rangle \in M : \mathcal{M} \models \neg\varphi(a_1, ..., a_n), \varphi \in \mathrm{Form}^n_{\mathcal{L}}\}.$$

- **Corollary**. (Inseparability of positive and negative fragments of ED) $\mathrm{ED}^{+}(\mathcal{M})$ *and* $\mathrm{ED}^{-}(\mathcal{M})$ *are definably inseparable in* $\mathcal{M}$, *i.e., there is no* $\mathcal{M}$*-definable* D *(parameters allowed) such that* $\mathrm{ED}^{+}(\mathcal{M}) \subseteq D$ *and* $\mathrm{ED}^{-}(\mathcal{M}) \cap D = \varnothing$.

Role of parameters

If $\mathcal{L}$ is a finite language, then it is possible for $\mathrm{Th}(\mathcal{M})$ to be definable within an $\mathcal{L}$-structure $\mathcal{M}$ if parameters are allowed.

- **Example 1.** $\mathrm{Th}(\mathcal{M})$ is parametrically definable in every recursively saturated model $\mathcal{M}$ of arithmetic or set theory. To see this, consider the recursive type $p(x)$ consisting of biconditionals of the form $\varphi \leftrightarrow \ulcorner\varphi\urcorner \in x$, where φ ranges over the recursive list of sentences, and the mapping $\varphi \mapsto \ulcorner\varphi\urcorner \in \omega$ is recursive.

- **Example 2.** Let $\mathcal{M}$ be a well-founded model of ZF of uncountable cofinality. Using the reflection theorem and the elementary chain theorem we can show that there is some ordinal α of $\mathcal{M}$ such that:

$$V_\alpha^{\mathcal{M}} \prec \mathcal{M}.$$

  By Tarski's definability of truth, $\mathrm{Th}(V_\alpha^{\mathcal{M}})$ is definable in $\mathcal{M}$ using the parameter α.

Syntactic formulation of TUT

- Let T be an $\mathcal{L}$-theory, and suppose $n \mapsto \mathbb{n}$ be an arbitrary mapping of ω (natural numbers) into the set of closed $\mathcal{L}$-terms (terms with no free variables).
- Fix an arbitrary 1-1 correspondence $\varphi \mapsto \#(\varphi)$ between $\mathrm{Form}_{\mathcal{L}}^{\leq 1}$ and ω, and let $n \longmapsto \varphi_n$ be its inverse.
- The *diagonal function* $\delta : \omega \to \omega$ is given by

$$\delta(n) = \#(\varphi_n(\mathbb{n})).$$

- A function $f : \omega \to \omega$ is said to be *T-definable* if there is an $\mathcal{L}$-formula $\theta(x, y)$ such that

$$\forall n \in \omega \;\; T \vdash \forall y \;\; [\theta(\mathbb{n}, y) \leftrightarrow y = \mathbb{f}(\mathbb{n})] \,.$$

- A subset P of ω is said to be *T-definable* if there is an $\mathcal{L}$-formula $\psi(x)$ such that:

$$\forall n \in P \quad T \vdash \psi(\mathbb{n}) \qquad \textbf{and} \qquad \forall n \notin P \quad T \vdash \neg\psi(\mathbb{n}).$$

Three relevant subtheories of PA

R1. $\vdash \underline{m} + \underline{n} = \underline{m+n}$
R2. $\vdash \underline{m} \cdot \underline{n} = \underline{m \cdot n}$
R3. $\vdash \underline{m} \neq \underline{n}, \quad \text{for } m \neq n$
R4. $\vdash x \leq \underline{n} \to \bigvee_{i \leq n} x = \underline{i}$
R5. $\vdash x \leq \underline{n} \vee \underline{n} \leq x$

Q1. $\vdash \mathsf{S}x = \mathsf{S}y \to x = y$
Q2. $\vdash \mathsf{S}x \neq 0$
Q3. $\vdash x = 0 \vee \exists y \ \ x = \mathsf{S}y$
Q4. $\vdash x + 0 = x$
Q5. $\vdash x + \mathsf{S}y = \mathsf{S}(x + y)$
Q6. $\vdash x \cdot 0 = 0$
Q7. $\vdash x \cdot \mathsf{S}y = x \cdot y + x$

PA$^-$1. $\vdash x + 0 = x$
PA$^-$2. $\vdash x + y = y + x$
PA$^-$3. $\vdash (x + y) + z = x + (y + z)$
PA$^-$4. $\vdash x \cdot 1 = x$
PA$^-$5. $\vdash x \cdot y = y \cdot x$
PA$^-$6. $\vdash (x \cdot y) \cdot z = x \cdot (y \cdot z)$
PA$^-$7. $\vdash x \cdot (y + z) = x \cdot y + x \cdot z$
PA$^-$8. $\vdash x \leq y \vee y \leq x$
PA$^-$9. $\vdash (x \leq y \wedge y \leq z) \to x \leq z$
PA$^-$10. $\vdash x + 1 \not\leq x$
PA$^-$11. $\vdash x \leq y \to (x = y \vee x + 1 \leq y)$
PA$^-$12. $\vdash x \leq y \to x + z \leq y + z$
PA$^-$13. $\vdash x \leq y \to x \cdot z \leq y \cdot z$
PA$^-$14. $\vdash x \leq y \to \exists z \ \ x + z = y$

Consequences of the syntactic formulation of TUT

- **Theorem 1.** (Syntactic formulation of TUT, ver 1.)
  Given a theory T, let $V_T = \{\#(\varphi) : T \vdash \varphi\}$. Assuming that T is consistent, then the diagonal function δ and the set V_T are not both T-definable.

- **Corollary.** (Syntactic formulation of TUT, ver. 2).
  If T is a consistent $\mathcal{L}$-theory such that δ is T-definable, then there is no $\mathcal{L}$-formula $\theta(x)$ such that for all $\mathcal{L}$-sentences φ we have:

$$T \vdash \varphi \leftrightarrow \theta(\mathbb{n}), \text{ where } n = \#(\varphi).$$

- **Corollary.** *If T is a consistent theory such that all total recursive functions are T-representable, and $\varphi \mapsto \#(\varphi)$ is recursive, then V_T is not recursive. In particular, T is incomplete.*

- **Remark.** *If T interprets Robinson's* R *(let alone Robinson's* Q*) , then all total recursive functions are T-definable.*

Proof of version 1 of TUT

Suppose not, thus there are formulae θ and ψ such that the following hold:

(1) $\forall n \in \omega \quad T \vdash \forall y\,[\theta(\mathbb{n}, y) \leftrightarrow y = \mathbb{r}]$, where $\delta(n) = r$.
(2) $\forall n \in V_T \quad T \vdash \psi(\mathbb{n})$.

(3) $\forall n \notin V_T \quad T \vdash \neg\psi(\mathbb{n})$.

Choose $m \in \omega$ such that $\varphi_m(x) = \forall y\,(\theta(x, y) \to \neg\psi(y))$, hence:

(4) $\varphi_m(\mathbb{m}) = \forall y\ \ (\theta(\mathbb{m}, y) \to \neg\psi(y))$.

If $T \vdash \varphi_m(\mathbb{m})$, then by (1) and (4) we have $T \vdash \neg\psi(\mathbb{k})$, where $\delta(m) = k$.

If $T \nvdash \varphi_m(\mathbb{m})$, then $\#(\varphi_m(\mathbb{m})) \notin V_T$; and by the definition of δ,

(5) $\delta(m) = \#(\varphi_m(\mathbb{m}))$.

So by (3) in this case we can also conclude that $T \vdash \neg\psi(\mathbb{k})$. Thus we have shown:

(6) $T \vdash \neg\psi(\mathbb{k})$.

By (1) and (6), $T \vdash \forall y\,(\theta(\mathbb{m}, y) \to \neg\psi(y))$. So by (4) and (5) $\delta(m) \in V_T$ and therefore by (2),

(7) $T \vdash \psi(\mathbb{k})$.

This contradicts the assumption of consistency of T.

Tarski's 1953 abstract

418*t*. Alfred Tarski: *Two general theorems on undefinability and undecidability*.

This paper contains a generalization of ideas known from works of Gödel and other authors. See specifically Mostowzki, *Sentences undeciable* $\cdots$, Amsterdam, 1952; R. M. Robinson, Proceedings of the International Congress of Mathematicians, 1950, vol. 1; Tarski, Studia Philosophica vol. 1. Assumptions: $\mathfrak{T}$ is any formalized theory; S is a set of $\mathfrak{T}$-formulas including all axioms of predicate calculus with identity and closed under rules of inference; $\Delta_0, \Delta_1, \cdots, \Delta_n, \cdots$ are $\mathfrak{T}$-terms containing no variables; $\sim(\Delta_0 = \Delta_1)$ is in S; x, y are fixed $\mathfrak{T}$-variables. A function F on and to the integers is called S-definable (relative to Δ_n) if, for some formula Φ and every integer n, $(x = \Delta_n) \rightarrow [(y = \Delta_{F(n)} \leftrightarrow \Phi]$ is in S. Consider a quite arbitrary one-one correlation between $\mathfrak{T}$-expressions Ψ and integers n; $Nr\ (\Psi)$ is the integer correlated with Ψ, Ω_n is the expression correlated with n. Let $D(n) \equiv Nr[(x = \Delta_n) \rightarrow \Omega_n]$; let $P(n)$ be 0 if Ω_n is in S, and 1 otherwise. Theorem I. *If S is consistent, then functions D and P are not both S-definable.* New assumptions: $G(n) \equiv Nr(\Delta_n)$ and $H(n, p) \equiv Nr(\Omega_n \frown \Omega_p)$ (where $\frown$ is the concatenation symbol) are general recursive functions. Then Theorem I implies: Theorem II. *If all general recursive functions are S-definable, then S is inconsistent or essentially undecidable.* (Received January 16, 1953.)

Tarski's assessment

The idea of this reconstruction and the realization of its far-reaching implications is due to Gödel [7]. The present version of this reconstruction is distinguished by its generality and simplicity. It applies to arbitrary formalized theories, and not only to those in which a comprehensive fragment of the arithmetic of natural numbers can be developed; to a large extent it is independent of the way in which the notion of validity has been defined for a given theory, and in particular it does not involve the notion of a formal proof within this theory; it does not use the apparatus of recursive functions—although this apparatus will play a fundamental role in applications of Theorem 1 to the decision problem.

For discussion on the history of TUT, see:

- R. Murawski, *Undefinability of truth; the problem of priority: Tarski vs Gödel*, **History and Philosophy of Logic** (1988), Vol. 19, 153–160.
- J. Woleński, *Gödel, Tarski and Truth*, **Revue Internationale de Philosophie** (2005) Vol. 59, pp. 459–490.

The Σ_n-hierarchy of arithmetical formulas (1)

Definition. Assuming t is an $\mathscr{L}_\mathrm{A}$-term not involving the variables $\bar{x}$, we abbreviate

$$\begin{array}{lccc} & \forall \bar{x}\ (\bar{x} < t \rightarrow \cdots) & \text{as} & \forall \bar{x}{<}t\ (\cdots) \\ \text{and} & \exists \bar{x}\ (\bar{x} < t \wedge \cdots) & \text{as} & \exists \bar{x}{<}t\ (\cdots), \end{array}$$

where $x_1, x_2, \ldots, x_k < t$ means $\bigwedge_{i=1}^{k} x_i < t$. Such quantifiers are said to be *bounded.* A formula is Δ_0 if all its quantifiers are bounded. Let $n \in \mathbb{N}$. A Σ_n*-formula* is one of the form

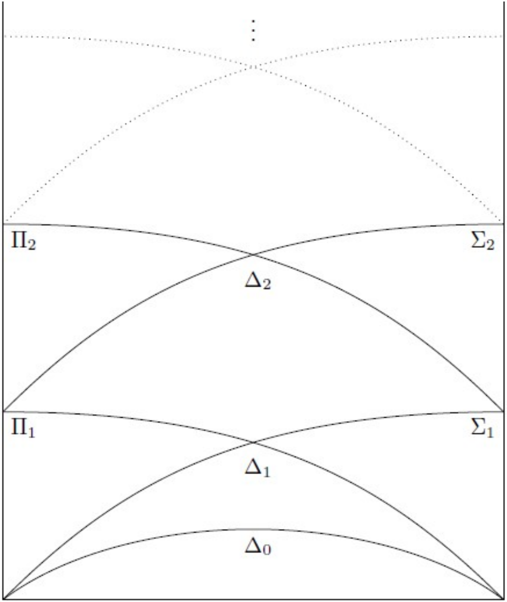
$$\exists \bar{x}_1\ \forall \bar{x}_2\ \cdots\ \mathrm{Q}\bar{x}_n\ \xi(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{z}),$$

where $\mathrm{Q} \in \{\forall, \exists\}$ and $\xi \in \Delta_0$. A Π_n*-formula* is one of the form

$$\forall \bar{x}_1\ \exists \bar{x}_2\ \cdots\ \mathrm{Q}'\bar{x}_n\ \zeta(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \bar{z}),$$

where $\mathrm{Q}' \in \{\forall, \exists\}$ and $\zeta \in \Delta_0$. Formulas that are equivalent to both a Σ_n- and a Π_n-formula are called Δ_n.

The Σ_n-hierarchy of arithmetical formulas (2)



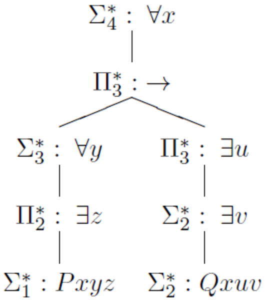
The Σ_n^*-hierarchy of formulas (1)

The Σ_n^* hierarchy measures the depth of quantifier alteration.

- $\Sigma_0^* := \Pi_0^* := \emptyset$.
- $\Sigma_{n+1}^* ::=$
  $\mathsf{AT} \mid \neg\, \Pi_{n+1}^* \mid (\Sigma_{n+1}^* \wedge \Sigma_{n+1}^*) \mid (\Sigma_{n+1}^* \vee \Sigma_{n+1}^*) \mid (\Pi_{n+1}^* \to \Sigma_{n+1}^*) \mid \exists v\, \Sigma_{n+1}^* \mid \forall v\, \Pi_n^*$.
- $\Pi_{n+1}^* ::=$
  $\mathsf{AT} \mid \neg\, \Sigma_{n+1}^* \mid (\Pi_{n+1}^* \wedge \Pi_{n+1}^*) \mid (\Pi_{n+1}^* \vee \Pi_{n+1}^*) \mid (\Sigma_{n+1}^* \to \Pi_{n+1}^*) \mid \forall v\, \Pi_{n+1}^* \mid \exists v\, \Sigma_n^*$.

The Σ_n^*-hierarchy of formulas (2)

Here is the parse-tree of $\forall x\,(\forall y\,\exists z\,Pxyz \to \exists u\,\exists v\,Qxuv)$ as an element of Σ_4^*.



MRDP (1)

- Hilbert's 10th problem (1900, among 23) was to find a general algorithm for deciding, given any n and any polynomial $f \in \mathbb{Z}[x_1, ..., x_n]$, whether or not f has a zero in $\mathbb{Z}^n$.

- **MRDP Theorem** (Matiyasevich, Robinson, Davis, Putnam, 1970).
  *A subset A of $\mathbb{N}$ is **recursively/computably enumerable** if and only if A is **Diophantine**, i.e., there exist polynomials $p(x, \overrightarrow{y}), q(x, \overrightarrow{y}) \in \mathbb{N}[x, \overrightarrow{y}]$ such that:*
  $$n \in A \iff \exists \overrightarrow{y} \in \mathbb{N} \;\; p(n, \overrightarrow{y}) = q(n, \overrightarrow{y}).$$

- Consequently, the existential theories of $\mathbb{N}$ and $\mathbb{Z}$ are undecidable, so Hilbert's problem has a negative answer.

MRDP (2)

- The MRDP theorem is provable in the fragment $\mathrm{I}\Delta_0 + \mathsf{Exp}$ of PA (Gaifman-Dimitracopoulos, 1980).
- By the MRDP-theorem each Σ_n-formula is equivalent to a Σ_n^*-formula in $\mathrm{I}\Delta_0 + \mathsf{Exp}$.
- So for arithmetical theories that extend $\mathrm{I}\Delta_0 + \mathsf{Exp}$ the Σ_n hierarchy and the Σ_n^* hierarchy coincide.

A general question (1)

- **General Question (GQ).** Suppose T_0 is a consistent theory. Does T_0 have a completion T that there is some $n \in \omega$ and some $A \subseteq \Sigma_n^*$ such that A axiomatizes T?
- Let $\mathbb{N}$ be the standard model of PA. For $T = \mathsf{TA} = \mathrm{Th}(\mathbb{N})$ the answer to GQ is in the negative. This follows from the Arithmetical Hierarchy Theorem of Kleene (1943) and Mostowski (1946) that states that $\Sigma_n^{\mathbb{N}} \subsetneq \Sigma_{n+1}^{\mathbb{N}}$ for each $n \in \omega$.
- For $T = \mathsf{PA}$, the negative answer follows from a theorem of Rabin (1961) that states that for for each $n \in \omega$ no consistent extension of PA (in the same language) is axiomatized by a set of Σ_n-sentences.
- Rabin's result refines an earlier theorem of Ryll-Nardzewski (1952) that states that no consistent extension of PA is finitely axiomatizable. Ryll-Nardzewski and Rabin both employed model-theoretic arguments relying on **nonstandard** elements to prove the aforementioned results.

Model theoretic proof of Rabin's Theorem (1)

- Given a model of arithmetic $\mathcal{M}$, $\mathrm{Th}_{\Pi_n}(\mathcal{M}) = \{\varphi \in \Pi_n : \mathcal{M} \models \varphi\}$.
- $\mathrm{PA}_{\Pi_n} = \{\varphi \in \Pi_n : \mathrm{PA} \vdash \varphi\}$.
- $\mathrm{I}\Sigma_n \subsetneq \mathrm{PA}_{\Pi_{n+2}}$.
- For $\mathcal{M} \models \mathrm{PA}$ and $n \in \omega$, $K_n(\mathcal{M})$ is the submodel of $\mathcal{M}$ whose universe consists of elements of $\mathcal{M}$ that are definable in $\mathcal{M}$ by a Σ_n-formula.
- $\mathcal{M} \prec_{\Pi_n} \mathcal{N}$ means that Π_n-formulae are absolute in the passage between the arithmetical structures $\mathcal{M}$ and $\mathcal{N}$.
- $\mathrm{B}\Sigma_n$ consists of universal generalizations of formulae of the following form, where ψ is a Σ_n-formula:

$$(\forall x \in a\ \exists y\ \psi(x,y)) \to (\exists b\ \forall x \in a\ \exists y \in b\ \psi(x,y)),$$

where the parameters of ψ are suppressed.

Model theoretic proof of Rabin's Theorem (2)

Theorem (Kirby-Paris, Lessan 1978). Suppose $n \in \omega$, $n \geq 1$, and $\mathcal{M}$ is a nonstandard model of PA, then:

(a) $K_n(\mathcal{M}) \prec_{\Pi_n} \mathcal{M}$, hence $K_n(\mathcal{M}) \models \mathrm{Th}_{\Pi_{n+1}}(\mathcal{M})$.

(b) If $K_n(\mathcal{M})$ is nonstandard, then $K_n(\mathcal{M}) \models \mathsf{PA}_{\Pi_{n+1}} + \mathsf{I}\Sigma_{n-1} + \neg \mathsf{B}\Sigma_n$.

The proof of part (a) involves a book-keeping argument together with a trick to collapse two existential quantifiers into one.

The proof of part (b) involves the Σ_n-definability of Σ_n-satisfaction within models of PA.

Proof theoretic proof of Rabin's Theorem (1)

- **Theorem.** (Mostowski Reflection Theorem 1952) *For each* $n \in \omega$, $\mathrm{Con}(\mathrm{True}_{\Sigma_n})$ *is provable in* PA.
- Mostowski's Reflection Theorem can be proved using the fact proved by Gentzen 1935 that every FOL proof can be replaced with another proof (with the same assumption and conclusion) that has the subformula property.
- The proof of Mostowski's Reflection Theorem uses the full induction scheme of PA in the verification that, for each *n* separately, PA can verify that True_{Σ_n} is closed under proofs with the subformula property.
- As we will see, Rabin's theorem can be established by using Mostowski's Reflection Theorem together with Gödel's second incompleteness theorem. This second proof gives less information.

Proof theoretic proof of Rabin's Theorem (2)

Here is more detail on the second proof of Rabin's theorem:

- Suppose to the contrary that T is an extension of PA that is axiomatized by a collection of $A\ \Sigma_n$-sentences for some $n \in \omega$.
- We can find an arithmetical formula, denoted $\mathsf{Prf}_{\mathrm{True}_n}(\pi, x)$ such that for each *standard* $\mathcal{L}$-sentence ψ and *standard* π, and each model $\mathcal{M}$ of PA, we have:

  $(*)$ $\mathcal{M} \models \mathsf{Prf}_{\mathrm{True}_n}(\pi, \ulcorner \psi \urcorner)$ iff π is (a code for) a proof of ψ from

  $$\mathsf{True}_n^{\mathcal{M}} := \{\varphi : \mathcal{M} \models \mathsf{True}_n(\ulcorner \varphi \urcorner)\}.$$

  Let

  $$\tau(x) := \exists y \mathsf{Prf}_{\mathsf{True}_n}(y, x).$$

  We will arrive at a contradiction by verifying that for all arithmetical sentences ψ, we have: $T \vdash \psi \leftrightarrow \tau(\ulcorner \psi \urcorner)$ (thus contradicting TUT).
- The left-to-right direction follows from (*). The other direction follows from (*) and Mostowski's Reflection Theorem.

Telegraphic History of Definable Partial Truth Predicates:

- Turing, Post, Kleene, Mostowski (1940s) – $0^{(n)}$ is Σ_n-complete.
- Mostowski (1952) – PA supports a definable truth predicate for Σ_n-formulae.
- Montague (1961) – Every inductive sequential theory supports a definable truth predicate for Σ_n^*-formulae.
- Levy (1965) – ZF supports a definable truth predicate for Σ_n^{Levy}-formulae.
- Gaifman and Dimitracopoulos (1980) – $\mathrm{I}\Delta_0 + \mathsf{Exp}$, supports a definable truth predicate for Σ_n-formulae.
- Pudlák (1984, 1998) – Every sequential theory supports a definable truth predicate for Depth_n-formulae.
- Visser (1994, 2019) – Every sequential theory supports a definable truth predicate for Σ_n^*-formulae.

Montague's generalization

- As shown by Montague (1961) Rabin's result can be generalized from PA to all *inductive* sequential theories T in a finite language. In the setting of Montague's result the relevant hierarchy is based on quantifier-alternations-depth.
- Canonical examples of inductive sequential theories include all extensions of PA, Z (Zermelo set theory), Z_2 (second order arithmetic), and KM (Kelley-Morse theory of classes).
- **Theorem.** (Montague Reflection Theorem 1959) *Suppose T is an inductive sequential theory formulated in a finite language. For each $n \in \omega$, $\mathrm{Con}(\mathrm{True}_{\Sigma_n^*})$ is provable in T.*

Inductive sequential theories (1)

- At first approximation, a theory is sequential if it supports a modicum of coding machinery to handle finite sequences of all objects in the domain of discourse. Gödel (1931) used the Chinese Remainder Theorem to show that PA is sequential. Jeřábek (2012) showed that PA^- is sequential, and Visser (2008) showed that Q is not sequential.
- It is known that T is sequential iff T has a definitional extension to *Adjunctive Set Theory*. The original definition of sequentiality due to Pudlák is as follows: A theory T is sequential if there is a formula $N(x)$ (read as "x is a number") , together with appropriate formulae providing interpretations of equality, and the operations of successor, addition, and multiplication for elements satisfying $N(x)$ such that T proves the translations of the axioms of Q when relativized to $N(x)$; and additionally, there is a formula $\beta(x, i, w)$ (whose intended meaning is that x is the i-th element of a sequence w) such that T proves that every sequence can be extended by any given element of the domain of discourse.

Inductive sequential theories (2)

The sequentiality axiom:

$$\forall w, x, k\, \exists w'\, \forall i, y \left[[N(k) \wedge i \leq k] \rightarrow \left[\begin{array}{c} \beta(y, i, w') \leftrightarrow \\ [i < k \wedge \beta(y, i, w)] \vee [i = k \wedge y = x] \end{array} \right] \right].$$

An inductive sequential theory T is a sequential theory within the full scheme of induction over N is provable. This is equivalent to saying that for all $\mathcal{M} \models T$, and any nonempty parameterically definable subset D of $\mathcal{M}$, $D \cap N^{\mathcal{M}}$ has a least element.

A Question of Lempp and Rossegger

- PA^- is the finitely axiomatized fragment of PA whose axioms describe the non-negative substructure of discretely ordered rings (with no instance of the induction scheme, hence the minus superscript).
- **Question.** *Is there a consistent completion of* $T = \mathrm{PA}^-$ *that is axiomatized by a set of sentences of bounded quantifier complexity*?
- The above question was posed by Steffen Lempp and Dino Rossegger in the context of their recent joint work [AGLRZ] with Uri Andrews, David Gonzalez, and Hongyu Zhu, in which they establish:

  Theorem. *The following are equivalent for a complete first-order theory* T:

  (1) *The set of models of* T *is* Π^0_ω*-complete under Wadge reducibility (i.e., reducibility via continuous functions).*

  (2) T *does not admit a first-order axiomatization by formulae of bounded quantifier complexity.*

  [AGLRZ] *The Borel complexity of the class of models of first-order theories*, arXiv:2402.10029[math.LO]

Preliminaries to "Fact F"

- The following result (FACT F) was established by Visser (1993, 2019); this result refines the work of Pudlák (1984, 1998) in which logical depth (length of the longest branch in the formation tree of the formula) is used as a measure of complexity instead of the depth of quantifier alternations complexity.

- The following conventions are at work in the statement of part (a) of Fact F for a given fixed interpretation $\mathcal{N}$ of Q in T.

  (1) An $\mathcal{L}$-formula $I = I(x)$ is a *T-provable definable cut* if T proves that the set of objects satisfying $I(x)$ is an initial segment of $\mathcal{N}$-numbers that satisfies $\mathrm{I}\Delta_0 + \Omega_1$, where Ω_1 is the axiom that states that the function $x \mapsto x^{\lfloor \log_2 x \rfloor}$ is a total function.[1]

  (2) Given a T-provable definable cut $I = I(x)$, the expression "x is a *Σ_n^*-formula in I*" is the conjunction of $\sigma_n(x)$ and $I(x)$, where $\sigma_n(x)$ is a designated $\mathcal{L}$-formula that expresses "x is the code of a Σ_n^*-formula".

[1] The treatment of syntax can be carried out fully within such a cut I, for example I is closed under conjunction of formulae in I. It is well-known that Q has a definable cut that satisfies $\mathrm{I}\Delta_0 + \Omega_1$, see Theorem 5.7 of [HP]. Thus a T-provable definable cut as defined here exists in every sequential theory.

Fact F

Fact F. *Suppose T is a sequential theory T formulated in a finite language $\mathcal{L}$, and $n \in \omega$. Fix some interpretation $\mathcal{N}$ of* Q *in T.*

(a) *There is a T-provable definable cut I_n* of $\mathcal{N}$ *and a formula* $\mathsf{Sat}_n(x, y)$ *such that, provably in T,* Sat_n *satisfies the Tarskian compositional clauses whenever x is a Σ_n^*-formula in I_n and for all variable assignments y.*

(b) *There is a formula* $\mathsf{True}_n(x)$ *such that, provably in T,* $\mathsf{True}_n(x)$ *is extensional*[2], *i.e., it respects the equivalence relation representing equality in the interpretation $\mathcal{N}$; and for all models $\mathcal{M} \models T$, and for all Σ_n^*-sentences ψ, we have:*

$$\mathcal{M} \models (\psi \leftrightarrow \mathsf{True}_n(\ulcorner \psi \urcorner)).$$

[2] Without this extensionality stipulation, the numeral does not necessarily work as a term.

Theorem A

Theorem A. *For any fixed $n \in \omega$, every consistent sequential theory formulated in a finite language that is axiomatized by a set of Σ_n^*-sentences is incomplete.*

Proof of Theorem A. Suppose not, and let T be consistent completion of sequential theory formulated in a finite language $\mathcal{L}$. Then by the definition of sequentiality T is also sequential. Suppose to the contrary that for some $n \in \omega$, T is axiomatized by a set of Σ_n^* sentences, i.e., suppose (1) below:

(1) For $n \in \omega$, there is a set A of Σ_n^* sentences such that for all $\mathcal{L}$-sentences ψ, $\psi \in T$ iff $A \vdash \psi$.

Our proof by contradiction of Theorem A will be complete once we verify Claim $\heartsuit$ below since it contradicts TUT.

CLAIM $\heartsuit$. There is a unary $\mathcal{L}$-formula $\varphi(x)$ such that for all $\mathcal{L}$-sentences ψ, $T \vdash \psi \leftrightarrow \varphi(\ulcorner \psi \urcorner)$.

Theorem A (cont'd)

Since T is sequential, we can find an $\mathcal{L}$-formula, denoted $\mathsf{Prf}_{\mathrm{True}_n}(\pi, x)$ such that for each *standard* $\mathcal{L}$-sentence ψ and *standard* π, and each model $\mathcal{M}$ of T, we have:

(2) $\mathcal{M} \models \mathsf{Prf}_{\mathrm{True}_n}(\pi, \ulcorner \psi \urcorner)$ iff π is (a code for) a proof of ψ from

$$\mathsf{True}_n^{\mathcal{M}} := \{\varphi : \mathcal{M} \models \mathsf{True}_n(\ulcorner \varphi \urcorner)\}.$$

Our proposed candidate of $\varphi(x)$ for establishing Claim $\heartsuit$ is the following formula $\rho(x)$; our choice of the letter ρ indicates the fact that the formula expresses *Rosser-provability* (from the true Σ_n^* sentences).

$$\rho(x) := \exists y \left[\mathsf{Prf}_{\mathsf{True}_n}(y, x) \wedge \forall z < y\ \neg \mathsf{Prf}_{\mathsf{True}_n}(z, \neg x)\right].$$

Thus our goal is to show that for all $\mathcal{L}$-sentences ψ, $T \vdash \psi \leftrightarrow \rho(\ulcorner \psi \urcorner)$. It suffices to show that for each model $\mathcal{M}$ of T, $\mathcal{M} \models \psi \leftrightarrow \rho(\ulcorner \psi \urcorner)$. For the rest of the proof, let $\mathcal{M} \models T$. We will first show:

(3) For all $\mathcal{L}$-sentences ψ, $\mathcal{M} \models \psi \rightarrow \rho(\ulcorner \psi \urcorner)$. To show (3), assume ψ holds in $\mathcal{M}$. Let n and A be as in (1), and note that $A \subseteq \mathsf{True}_n^{\mathcal{M}}$.

Theorem A (concluded)

By the assumptions about T, there are finitely many sentences $\alpha_1, ..., \alpha_n$ in A such that $\{\alpha_1, ..., \alpha_n\} \vdash \psi$. Let $\pi_0 \in \omega$ be (the code of) a proof of ψ from $\{\alpha_1, ..., \alpha_n\}$. Thanks to (2) we have: $\mathcal{M} \models \mathsf{Prf}_{\mathsf{True}_n}(\pi_0, \ulcorner\psi\urcorner)$. The assumption of consistency of T coupled with (2) yields: $\mathcal{M} \models \forall z < \pi_0\ \neg\mathsf{Prf}_{\mathsf{True}_n}(z, \ulcorner\neg\psi\urcorner)$. Hence (3) holds.

To complete the proof of CLAIM $\heartsuit$, we need to show that $\mathcal{M} \models \neg\psi \to \neg\rho(\ulcorner\psi\urcorner)$ for all $\mathcal{L}$-sentences ψ. For this purpose assume $\mathcal{M} \models \neg\psi$. By putting (1) and the assumption that $\mathcal{M} \models \neg\psi$, we conclude that there is a *standard* proof π_0 of $\neg\psi$ from $\mathsf{True}_n^{\mathcal{M}}$, which by (2) implies:

(4) For some $\pi_0 \in \omega$, $\mathcal{M} \models \mathsf{Prf}_{\mathsf{True}_n}(\pi_0, \ulcorner\neg\psi\urcorner)$.

To see that $\mathcal{M} \models \neg\rho(\ulcorner\psi\urcorner)$ suppose to the contrary that $\mathcal{M} \models \rho(\ulcorner\psi\urcorner)$. By the choice of ρ, this means:

(5) For some $m_0 \in M$, $\mathcal{M} \models \mathsf{Prf}_{\mathsf{True}_n}(m_0, \ulcorner\psi\urcorner) \wedge \forall z < m_0\ \neg\mathsf{Prf}_{\mathsf{True}_n}(z, \ulcorner\neg\psi\urcorner)$.

The key observation is that putting (2) with the assumption $\mathcal{M} \models \neg\psi$ allows us to conclude that the m_0 in (5) must be a **nonstandard element of** $\mathcal{M}$. Thus by standardness of π_0 of (4) and the ordering properties of 'natural numbers' in $\mathcal{M}$, $\mathcal{M} \models \pi_0 < m_0$, which contradicts the second conjunct of (5). $\square$

A more general form of Theorem A

Theorem $\mathbf{A}^{+}$. *Let T be a computably enumerable sequential theory formulated in a finite language $\mathcal{L}$ and suppose A is a collection of $\mathcal{L}$-sentences such that $A \subseteq \Sigma_n^*$ for some $n \in \omega$ and $T \cup A$ is consistent. Then $T \cup A$ is incomplete*.

Theorem B and B$^+$

Theorem B. *For each $n \in \omega$ every consistent extension of* $\mathrm{I}\Delta_0 + \mathrm{Exp}$ (*in the same language*) *that is axiomatized by a set of Σ_n-sentences is incomplete.*

Proof. As shown by Gaifman and Dimitracopoulos (1980) for each $n \in \omega$ there is a formula Sat_{Σ_n} such that, provably in $\mathrm{I}\Delta_0 + \mathrm{Exp}$, Sat_{Σ_n} satisfies compositional clauses for all Σ_n-formulae. In particular there is a formula $\mathrm{True}_{\Sigma_n}(x)$ such that for all models $\mathcal{M}$ of $\mathrm{I}\Delta_0 + \mathrm{Exp}$, and for all Σ_n-sentences ψ, $\psi \in \mathrm{True}^{\mathcal{M}}_{\Sigma_n}$ iff $\mathcal{M} \models \psi$. We can now repeat the proof strategy of Theorem A with the use of $\mathrm{True}^{\mathcal{M}}_{\Sigma_n}$ instead of $\mathrm{True}^{\mathcal{M}}_{n}$.

Alternatively, invoke the provability of the MRDP theorem in $\mathrm{I}\Delta_0 + \mathrm{Exp}$. By the MRDP-theorem each Σ_n-formula is equivalent to a Σ^*_n-formula in $\mathrm{I}\Delta_0 + \mathrm{Exp}$, so Theorem A applies. □

We can similarly obtain an analogous strengthening Theorem B$^+$ of Theorem B.

Theorem A$^+$. *Let T be a computably enumerable extension of* $\mathrm{I}\Delta_0 + \mathrm{Exp}$ (*in the same language*) *and suppose A is a collection of arithmetical sentences such that $A \subseteq \Sigma_n$ for some $n \in \omega$ and $T \cup A$ is consistent. Then $T \cup A$ is incomplete.*

Analogue for set theory (1)

The set-theoretical analogue of Theorem A is Theorem C below concerning the well-known Levy hierarchy of formulae of set theory, which is the set-theoretical counterpart of the Σ_n-hierarchy of arithmetical formulae. Theorem C can be proved with the same strategy used in the proof of Theorem A (and the first proof of Theorem B) thanks to the availability of the relevant definable partial satisfaction classes in KP.

Here KP is Kripke-Platek set theory with the scheme of foundation limited to Π_1^{Levy}-formulae (equivalently: the scheme of $\in$-induction for Σ_1^{Levy}-formulae). Thus in contrast to Barwise's KP in his book "Admissible Sets and Structures", which includes the full scheme of foundation, our version of KP is finitely axiomatizable. Note that the axiom of infinity is not among the axioms of KP.

The existence of definable partial satisfaction classes in KP follows from two facts: (1) KP can prove that every set is contained in a transitive set; and (2) KP can define the satisfaction predicate for all of its internal set structures (the proofs of both of these facts can be found in Barwise's book; the proofs therein make it clear that only Π_1^{Levy}-Foundation is invoked).

Analogue for set theory (2)

Theorem C. *For each $n \in \omega$ every consistent completion of* KP (*in the same language*) *that is axiomatized by a set of Σ_n^{Levy}-sentences is incomplete.*

Similar to Theorems A and B, Theorem C can be readily generalized to:

Theorem C$^+$. *Let T be a computably enumerable extension of* KP (*in the same language*) *and suppose A is a collection of set-theoretical sentences such that $A \subseteq \Sigma_n^{\mathrm{Levy}}$ for some $n \in \omega$ and $T \cup A$ is consistent. Then $T \cup A$ is incomplete.*

Remark. In the above theorems KP can be replaced by "Mac Lane set theory", which is Zermelo set theory with comprehension restricted to Δ_0-formulas together with the sentence "every set is contained in a transitive set".

Connected ideas

> **Proposition:** Let T_0 be an r.e. theory interpreting Robinson's arithmetic, and Γ a set of sentences for which T_0 has a truth predicate $\mathrm{Tr}_\Gamma(x)$, that is,
>
> $$T_0 \vdash \phi \leftrightarrow \mathrm{Tr}_\Gamma(\overline{\ulcorner \phi \urcorner}) \qquad (*)$$
>
> for all $\phi \in \Gamma$. Then no extension of T_0 by a set of Γ-sentences is a consistent complete theory.

SOURCE: A MATHOVERFLOW ANSWER BY Emil Jeřábek (2016);

https://mathoverflow.net/questions/256785/a-completion-of-zfc

- As shown by Mateusz Łełyk and Bartosz Wcisło in their recent paper *Universal properties of truth* Theorem A can also be established using the machinery of so-called (n, k)-flexible formulas.

Conceptual/Pedagogical take-away

Let R be the well-known fragment of PA within which all recursive functions are representable.

One can prove the (first) incompleteness theorem for a consistent computably enumerable extension T of R, without any extra soundness assumptions about T, by first proving the syntactic version of Tarski's undefinability of truth theorem, and then the incompleteness of T can be demonstrated using a reductio ad absurdum by verifying that the completeness of T implies that Rosser provability from T yields a truth definition. Technically, this falls under our Theorem A^{+}, by setting $A = \varnothing$ in that theorem.

Note that in contrast to the usual proof of the incompleteness theorem using the fixed point theorem, our proof is not constructive, i.e., it does not yield an algorithm that takes a description of a consistent computably enumerable extension T of R as input and outputs a sentence that is independent of T.

Why $\mathcal{L}$ cannot be infinite (1)

- Consider the theory $U = \mathsf{CT}^-_\omega[\mathrm{I}\Sigma_1]$ of ω-iterated compositional truth over $\mathrm{I}\Sigma_1$ (without any induction for formulae using nonarithmetical symbols, hence the minus superscript) formulated in an extension of the language $\mathcal{L}_\mathsf{A}$ of arithmetic with infinitely many predicates $\{\mathsf{T}_{n+1} : n \in \omega\}$, and Tarski-style compositional axioms that stipulate that T_{n+1} is compositional for all $\mathcal{L}_n$-formulae, with $\mathcal{L}_0 = \mathcal{L}_\mathsf{A}$ and $\mathcal{L}_{n+1} = \mathcal{L}_n \cup \{\mathsf{T}_{n+1}\}$.
- Since bi-conditionals of form $\varphi \longleftrightarrow \mathsf{T}_{n+1}(\ulcorner\varphi\urcorner)$ are provable in U for every $\mathcal{L}_n$-sentence (thanks to the available composition axioms) ANY complete extension V of U is axiomatized by U (which is of bounded complexity) together with atomic sentences of form $\mathsf{T}_{n+1}(\ulcorner\varphi\urcorner)$ where $\varphi \in V$ and φ is an $\mathcal{L}_n$-sentence, thus U axiomatizable by a set of axioms of bounded quantifier complexity.
- By adding one axiom (internal induction) to the above theory we can get a theory of bounded complexity whose deductive consequence includes PA, and every completion of which is boundely axiomatizable.

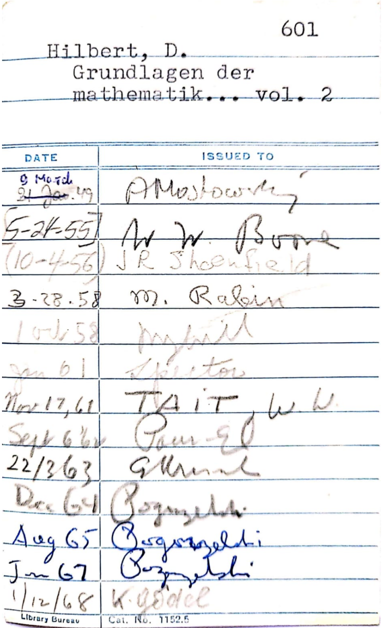
Why $\mathcal{L}$ cannot be infinite (2)

- Alternatively, starting with any theory T formulated in a language $\mathcal{L}$, we can apply a process known in model theory as *Morleyization/Atomization* to obtain an extension T^+ of T, formulated in an extension $\mathcal{L}^+$ of $\mathcal{L}$, such that T^+ is axiomatized by adding a collection of sentences of bounded quantifier depth to T, and T^+ has elimination of quantifiers in the sense that for each $\mathcal{L}^+$-formula $\varphi(x_1, ..., x_n)$, there is an n-ary predicate $P_\varphi \in \mathcal{L}^+$ such that the equivalence $\varphi(x_1, ..., x_n) \leftrightarrow P_\varphi(x_1, ..., x_n)$ is provable in T^+.

Some Off Shoots

- **Question.** *Is it possible for a consistent completion of* Q *to be axiomatized by a collection of sentences of bounded quantifier-depth*? Conjecture: Yes.
- Recall that Theorem B concerned the usual Σ_n hierarchy of arithmetical formulae and theories extending $\mathrm{I}\Delta_0 + \mathrm{Exp}$. A natural question is to explore the extent to which textrm$\mathrm{I}\Delta_0 + \mathrm{Exp}$ can be weakened.
- In joint work with Albert Visser and Mateusz Łełyk (forthcoming) we show the following result.
- **Theorem.** *There is a consistent completion of* PA^- *(in the same language) that is axiomatized by single sentence together with a set of* Σ_1*-sentences.*
- The result above can be extended to the stronger theory $\mathrm{PA}^- + \mathrm{Collection}$.
- Indeed, the technique can be pushed to even obtain a similar result for the theory $\mathrm{IOpen} + \mathrm{Collection}$.

Thanx!



601

Hilbert, D.
Grundlagen der
mathematik... vol. 2

DATE	ISSUED TO
9 March ~~21 Jan.~~ 49	A Mostowski
5-24-55	W. W. Boone
10-4-56	J R Shoenfield
3-28.58	M. Rabin
1 Oct 58	Myhill
Jan 61	Spector
Nov 17, 61	TAIT, W. W.
Sept 6 '62	Paul E[illegible]
22/3 63	G Kreisel
Dec 64	[illegible]
Aug 65	[illegible]
Jan 67	[illegible]
1/12/68	K. Gödel

Library Bureau Cat. No. 1152.5