Löb's Theorem and Provability Predicates in Coq

Final Bachelor Talk

Janis Bailitis

Saarland University
Advisors: Dr. Yannick Forster, Dr. Dominik Kirst
Supervisor: Prof. Dr. Gert Smolka
Programming Systems Lab

August 23, 2024

Introduction

- Sufficiently strong formal systems S have **provability predicates** $\mathrm{Pr}(x) : \mathbb{F}$
  - $\mathrm{Pr}(x)$ asserts provability of other formulas ($S \vdash \varphi$ iff $S \vdash \mathrm{Pr}(\overline{\varphi})$)
  - Many different of various strengths, even for same formal system

Theorem (Gödel, 1931)

If $\mathrm{Pr}(x)$ and S are sufficiently strong, and $S \vdash \varphi \leftrightarrow \neg\mathrm{Pr}(\overline{\varphi})$, then φ is independent.

Problem (Henkin, 1952)

What happens if $S \vdash \varphi \leftrightarrow \mathrm{Pr}(\overline{\varphi})$?

Theorem (Löb, 1955)

If $\mathrm{Pr}(x)$ and S are sufficiently strong, then Henkin's formulas are provable.

Löb's Theorem and Motivation

Theorem (Löb's theorem, 1955)

Let $\mathrm{Pr}(x)$ *and* S *be sufficiently strong. For all sentences* φ,

$$(S \vdash \mathrm{Pr}(\overline{\varphi}) \to \varphi) \textit{ implies } (S \vdash \varphi).$$

- Implies **Gödel's second incompleteness theorem**
  - Mechanised only once: Paulson (2015) in Isabelle. Tedious task
  - Paulson's proof easily extends to Löb's theorem
- **Gödel's first incompleteness theorem** mechanised often[1]
- Kirst and Peters: Computational proof of Gödel's first incompleteness theorem
- Leave Gödel's second incompleteness theorem as future work

This thesis: Is there a proof of Löb's theorem à la Kirst and Peters?

[1]Shankar (1986); O'Connor (2005); Harrison (2009); Paulson (2015); Popescu and Traytel (2019); Kirst and Peters (2023)

‘Sufficiently Strong’ in View of Löb’s Theorem

‘Sufficiently strong’ provability predicates satisfy the HBL conditions:

Hilbert-Bernays-Löb (HBL) Conditions (Hilbert-Bernays (1939), Löb (1955))

$\mathrm{Pr}(x) : \mathbb{F}$ satisfies

- **necessitation** if $S \vdash \varphi$ implies $S \vdash \mathrm{Pr}(\overline{\varphi})$
- **internal necessitation** if $S \vdash \mathrm{Pr}(\overline{\varphi}) \to \mathrm{Pr}(\overline{\mathrm{Pr}(\overline{\varphi})})$
- **the distributivity law** if $S \vdash \mathrm{Pr}(\overline{\varphi \to \psi}) \to \mathrm{Pr}(\overline{\varphi}) \to \mathrm{Pr}(\overline{\psi})$

‘Sufficiently strong’ theories S have the diagonalisation property:

Diagonalisation Property (Carnap (1934))

S has the **diagonalisation property** if for all $\varphi(x) : \mathbb{F}$ there is $G : \mathbb{F}$ s.t.

$$S \vdash G \leftrightarrow \varphi(\overline{G}).$$

HBL + Diagonalisation property = Löb’s theorem (20 lines Coq, 60 lines Isabelle)

Church's Thesis (CT)

- This thesis: Formal systems of first-order arithmetic
- CT: 'Every function is computable in a concrete model of computation.'[1]
- Results based on a variant of CT for arithmetic ($\mathsf{CT}_{\mathsf{PA}}$ / CT_{Q}):[2]

Axiom ($\mathsf{CT}_{\mathsf{PA}}$, Hermes and Kirst (2022))

For all $f : \mathbb{N} \to \mathbb{N}$ there is $\varphi_f(x_1, x_2) : \mathbb{F}$ such that for all $n : \mathbb{N}$,

$$\mathsf{PA} \vdash \forall y.\, \varphi_f(\overline{n}, y) \leftrightarrow y = \overline{f\,n}.$$

Lemma (Representability, cf. Hermes and Kirst (2022))

If $P : \mathbb{N} \to \mathbb{P}$ is enumerable, there is $\varphi_P(x) : \mathbb{F}$ such that $P\,n$ iff $\mathsf{PA} \vdash \varphi_P(\overline{n})$.

[1] Kreisel (1965) as well as Troelstra and van Dalen (1988).

[2] The thesis uses EPF_μ (Richman (1983), Forster (2021)) implying $\mathsf{CT}_{\mathsf{PA}}$ (Kirst and Peters (2023)).

Exploiting Church's Thesis

$\lambda\varphi.\,\mathsf{PA} \vdash \varphi$ is enumerable (mechanised by Forster, Kirst, and Smolka (2019))

Corollary

There is $\mathsf{Pr}_{\mathsf{CT}}(x) : \mathbb{F}$ *such that* $\mathsf{PA} \vdash \varphi$ *iff* $\mathsf{PA} \vdash \mathsf{Pr}_{\mathsf{CT}}(\overline{\varphi})$.

$\mathsf{CT}_{\mathsf{PA}}$ easily shows that PA has the diagonalisation property.

Lemma (Diagonal Lemma, Carnap (1934))

For all $\varphi(x) : \mathbb{F}$ *there is* $G : \mathbb{F}$ *s.t.* $\mathsf{PA} \vdash G \leftrightarrow \varphi(\overline{G})$.

- Gödel's first incompleteness theorem (1931), with Rosser's strengthening[1]
- Tarski's theorem (1935)
- Essential undecidability of PA

[1] Needs slight strengthening of $\mathsf{CT}_{\mathsf{PA}}$ which also follows from EPF_μ (Kirst and Peters (2023)).

External and Internal Provability

Is $\mathrm{Pr}_{\mathrm{CT}}(x)$ sufficiently strong for Löb's theorem? $\mathrm{Pr}_{\mathrm{CT}}(x)$ is external:

Definition (External Provability Predicates (Kreisel, 1953)[1])

$\mathrm{Pr}(x) : \mathbb{F}$ is **external provability predicate** for T if $T \vdash \varphi$ iff $T \vdash \mathrm{Pr}(\overline{\varphi})$.

Definition (Internal Provability Predicates)

$\mathrm{Pr}(x) : \mathbb{F}$ is **internal provability predicate** if it is both

- an external provability predicate, and
- satisfies the HBL conditions.

Does $\mathrm{Pr}_{\mathrm{CT}}(x)$ satisfy the HBL conditions?

[1] Kreisel did not introduce the terminology external / internal. Feferman (1960) first used such terms (extensional and intensional).

Church's Thesis and Löb's Theorem

Does $\mathrm{Pr}_{\mathrm{CT}}(x)$ satisfy the HBL conditions? Not necessarily!

Definition (Mostowski's Modification, 1965[1])

The **Mostowski modification** $\mathrm{Pr}^M(x) : \mathbb{F}$ of $\mathrm{Pr}(x)$ is

$$\mathrm{Pr}^M(x) := \mathrm{Pr}(x) \land x \neq \overline{\bot}.$$

$\mathrm{Pr}^M(x)$ does not satisfy distributivity law: $\mathrm{PA} \nvdash \mathrm{Pr}^M(\overline{\varphi \to \bot}) \to \mathrm{Pr}^M(\overline{\varphi}) \to \mathrm{Pr}^M(\overline{\bot})$

Lemma

$\mathrm{Pr}^M_{\mathrm{CT}}(x)$ *is an external provability predicate, too.*

Abstract perspective too weak!

[1] This particular formulation is from Bezboruah and Shepherdson (1976).

Defining an Internal Provability Predicate

Gödel's Approach (also widely used in literature)

- $\mathrm{Prf}(w, x) : \mathbb{F}$ checking that w is a proof of x (w is seen as list of formulas)
- $\mathrm{Pr}(x) := \exists w.\, \mathrm{Prf}(w, x)$

Hilbert System (à la Rautenberg (2010), Troelstra and Schwichtenberg (2000))

Let $\mathcal{H}$ be a finite set of formulas. $\mathrm{PA} \vdash_{\mathcal{H}} \varphi$ is defined inductively.

$$\frac{\mathrm{PA} \vdash_{\mathcal{H}} \varphi \to \psi \qquad \mathrm{PA} \vdash_{\mathcal{H}} \varphi}{\mathrm{PA} \vdash_{\mathcal{H}} \psi} \qquad \frac{\varphi \in \mathcal{H}}{\mathrm{PA} \vdash_{\mathcal{H}} \forall x_1. \ldots x_n. \varphi} \qquad \frac{\varphi \in \mathrm{PA}}{\mathrm{PA} \vdash_{\mathcal{H}} \varphi}$$

For right choice of $\mathcal{H}$: Have $\mathrm{PA} \vdash \varphi$ iff $\mathrm{PA} \vdash_{\mathcal{H}} \varphi$.

Defining an Internal Provability Predicate (Continued)

- Gödel's provability predicate uses list functions
- List functions not native to PA $\rightarrow$ tedious to define (see Boolos (1993))

Definition (Extended Signature of Peano Arithmetic, simplified)

In addition to the symbols of PA, EPA contains the following function symbols:

$[\,]$ (nil)	$\lvert\ell\rvert$ (length)	$\ell +\!\!\!+ \ell'$ (append)
$x :: \ell$ (cons)	$\ell[i]$ (indexed access)	$x \rightsquigarrow y$ (implication)

Based on such a definition, we

1. defined a candidate for an internal provability predicate, and
2. mechanised necessitation as well as the distributivity law for this candidate.

Contributions

Is there a proof of Löb's theorem à la Kirst and Peters? No!

- Mechanised proof of Löb's theorem
  - For first-order arithmetic in Coq assuming HBL conditions and $\mathrm{CT}_{\mathrm{PA}}$
  - For HF set theory in Isabelle based on Paulson's development
- Mechanised diagonal lemma and important limitative theorems assuming $\mathrm{CT}_{\mathrm{PA}}$
- Analysed why $\mathrm{CT}_{\mathrm{PA}}$ is too weak for Löb's theorem
- Mechanised extension of PA easing definition of internal provability predicates
- Gave candidate for internal provability predicate and parts of correctness proof

Future Work

- Mechanise internal necessitation
- Decide whether to keep using extended PA
- Contribute Isabelle development to Archive of Formal Proofs[1]
- Contribute Coq development to Coq Library of First-Order Logic [Kir+22]
- Mechanise axiom-free proof of diagonal lemma and limitative theorems

Discussion.

[1] https://www.isa-afp.org/

[Boo93] George S. Boolos. **The Logic of Provability.** 5th. Cambridge University Press, 1993.

[BS76] A. Bezboruah and John C. Shepherdson. **'Gödel's Second Incompleteness Theorem for Q'.** In: **The Journal of Symbolic Logic** 41.2 (1976), pp. 503–512. DOI: 10.1017/S0022481200051586.

[Car34] Rudolf Carnap. **Logische Syntax der Sprache.** 1st. Schriften zur wissenschaftlichen Weltauffassung. Springer Berlin, Heidelberg, 1934.

[DT88] Dirk van Dalen and Anne S. Troelstra. **Constructivism in Mathematics. An Introduction.** Elsevier Science Publishers B.V., 1988. ISBN: 0-444-70266-0.

References ii

[Fef60] Solomon Feferman. **'Arithmetization of metamathematics in a general setting'.** In: **Fundamenta mathematicae** 49 (1 1960), pp. 35–92.

[FKS19] Yannick Forster, Dominik Kirst and Gert Smolka. **'On Synthetic Undecidability in Coq, with an Application to the Entscheidungsproblem'.** In: **Proceedings of the 8th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2019, Cascais, Portugal, January 14-15, 2019**. Ed. by Assia Mahboubi and Magnus O. Myreen. ACM, 2019, pp. 38–51. DOI: 10.1145/3293880.3294091.

[For21] Yannick Forster. **'Church's Thesis and Related Axioms in Coq's Type Theory'.** In: **29th EACSL Annual Conference on Computer Science Logic, CSL 2021, January 25-28, 2021, Ljubljana, Slovenia (Virtual Conference)**. Ed. by Christel Baier and Jean Goubault-Larrecq. Vol. 183. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, 21:1–21:19. DOI: 10.4230/LIPICS.CSL.2021.21.

[Gö31] Kurt Gödel. **'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I'.** In: **Monatshefte für Mathematik** 38.1 (1931), pp. 173–198.

[Har09] John Harrison. **Handbook of Practical Logic and Automated Reasoning.** Cambridge University Press, 2009. DOI: 10.1017/CBO9780511576430.

[HB39] David Hilbert and Paul Bernays. **Grundlagen der Mathematik.** 1st. Vol. 2. Berlin: Springer, 1939.

[Hen52] Leon Henkin. **'A problem concerning provability'.** In: **The Journal of Symbolic Logic** 17.2 (1952), p. 160. ISSN: 00224812. URL: http://www.jstor.org/stable/2266288.

[HK22] Marc Hermes and Dominik Kirst. **'An Analysis of Tennenbaum's Theorem in Constructive Type Theory'.** In: **7th International Conference on Formal Structures for Computation and Deduction (FSCD 2022)**. Ed. by Amy P. Felty. Vol. 228. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 9:1–9:19. ISBN: 978-3-95977-233-4. DOI: 10.4230/LIPIcs.FSCD.2022.9.

References v

[HK23] Marc Hermes and Dominik Kirst. **‘An Analysis of Tennenbaum’s Theorem in Constructive Type Theory’.** In: **CoRR** abs/2302.14699 (2023). DOI: 10.48550/ARXIV.2302.14699. arXiv: 2302.14699.

[Kir+22] Dominik Kirst et al. **‘A Coq Library for Mechanised First-Order Logic’.** In: **The Coq Workshop** (2022).

[KP23] Dominik Kirst and Benjamin Peters. **‘Gödel’s Theorem Without Tears - Essential Incompleteness in Synthetic Computaility’.** In: **31st EACSL Annual Conference on Computer Science Logic (CSL 2023)**. Ed. by Bartek Klin and Elaine Pimentel. Vol. 252. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 30:1–30:18. ISBN: 978-3-95977-264-8. DOI: 10.4230/LIPIcs.CSL.2023.30.

[Kre53] Georg Kreisel. **'On a problem of Henkin's'.** In: **Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings**. A 56 (1953), pp. 405–406.

[Kre65] Georg Kreisel. **'Mathematical logic'.** In: **Lectures on Modern Mathematics** 3 (1965), pp. 95–195.

[Lö55] Martin H. Löb. **'Solution of a Problem of Leon Henkin'.** In: **The Journal of Symbolic Logic** 20.2 (1955), pp. 115–118. DOI: 10.2307/2266895.

[Mos65] Andrzej Mostowski. **'Thirty years of foundational studies: lectures on the development of mathematical logic and the study of the foundations of mathematics in 1930-1964'.** In: **Acta Philosophica Fennica** 17 (1965), pp. 1–180.

[O'C05] Russell O'Connor. **'Essential Incompleteness of Arithmetic Verified by Coq'.** In: **Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005, Oxford, UK, August 22-25, 2005, Proceedings**. Ed. by Joe Hurd and Thomas F. Melham. Vol. 3603. Lecture Notes in Computer Science. Springer, 2005, pp. 245–260. DOI: 10.1007/11541868_16.

[Pau15] Lawrence C. Paulson. **'A Mechanised Proof of Gödel's Incompleteness Theorems Using Nominal Isabelle'.** In: **Journal of Automated Reasoning** 55.2 (2015), pp. 1–37. DOI: 10.1007/s10817-015-9322-8.

[PT21] Andrei Popescu and Dmitriy Traytel. **'Distilling the Requirements of Gödel's Incompleteness Theorems with a Proof Assistant'.** In: **Journal of Automated Reasoning** 65.7 (2021), pp. 1027–1070. DOI: 10.1007/S10817-021-09599-8.

[Rau10] Wolfgang Rautenberg. **A Concise Introduction to Mathematical Logic.** 3rd. Springer New York Dordrecht Heidelberg London, 2010. DOI: 10.1007/978-1-4419-1221-3.

[Ric83] Fred Richman. **'Church's Thesis Without Tears'.** In: **The Journal of Symbolic Logic** 48.3 (1983), pp. 797–803. DOI: 10.2307/2273473.

[Ros36] J. Barkley Rosser. **'Extensions of Some Theorems of Gödel and Church'.** In: **The Journal of Symbolic Logic** 1.3 (1936), pp. 87–91. DOI: 10.2307/2269028.

[Sha86] Natarajan Shankar. **'Proof-checking metamathematics'.** PhD thesis. University of Texas, 1986.

[Sha94] Natarajan Shankar. **Metamathematics, Machines and Gödel's Proof.** Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1994. DOI: 10.1017/CBO9780511569883.

[Smo24] Gert Smolka. **Modeling and Proving in Computational Type Theory Using the Coq Proof Assistant.** Textbook under construction. (visited on 05 July 2024). 2024. URL: https://www.ps.uni-saarland.de/~smolka/drafts/mpctt.pdf.

[Tar35] Alfred Tarski. **'Der Wahrheitsbegriff in den formalisierten Sprachen'.** In: **Studia Philosophica. Commentarii Societatis Philosophicae Polonorum** 1 (1935), pp. 261 –405. URL: https://www.sbc.org.pl/dlibra/publication/24411/edition/21615.

[TS00] Anne S. Troelstra and Helmut Schwichtenberg. **Basic Proof Theory.** 2nd. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2000. DOI: 10.1017/CBO9781139168717.

Mechanisation

Coq

- 2600 lines of code (600 specification, 1900 proof, 100 comment)
- Most intricate proof: Distributivity law in EHA (about 400 lines of code)
- Lots of code dealing with substitutions

Isabelle

- 100 lines of code (60 for Löb proof, 40 for lemmas)
- Can still be shortened

Background: Used Hilbert System

$\mathcal{H}(\varphi \to \psi \to \varphi)$ $\quad$ $\mathcal{H}((\varphi \to \psi \to \tau) \to (\psi \to \tau) \to \varphi \to \tau)$

$\mathcal{H}(\varphi \to \psi \to \varphi \wedge \psi)$ $\quad$ $\mathcal{H}(\varphi \wedge \psi \to \varphi)$

$\mathcal{H}(\varphi \to \varphi \vee \psi)$ $\quad$ $\mathcal{H}(\varphi \wedge \psi \to \psi)$

$\mathcal{H}(\psi \to \varphi \vee \psi)$ $\quad$ $\mathcal{H}(\varphi \vee \psi \to (\varphi \to \tau) \to (\psi \to \tau) \to \tau)$

$\mathcal{H}(\bot \to \varphi)$ $\quad$ $\mathcal{H}(\varphi \to \forall x.\, \varphi)$ $\quad$ x fresh for φ

$\mathcal{H}((\forall x.\, \varphi) \to \varphi[x \mapsto t])$ $\quad$ $\mathcal{H}((\forall x.\, \varphi \to \psi) \to (\forall x.\, \varphi) \to \forall x.\, \psi)$

$\mathcal{H}(\varphi[x \mapsto t] \to \exists x.\, \varphi)$ $\quad$ $\mathcal{H}((\exists x.\, \varphi) \to (\forall x.\, \varphi \to \psi) \to \psi)$ $\quad$ x fresh for ψ

$$\frac{\mathrm{PA} \vdash_{\mathcal{H}} \varphi \to \psi \qquad \mathrm{PA} \vdash_{\mathcal{H}} \varphi}{\mathrm{PA} \vdash_{\mathcal{H}} \psi} \qquad \frac{\varphi \in \mathcal{H}}{\mathrm{PA} \vdash_{\mathcal{H}} \forall x_1. \ldots x_n.\, \varphi} \qquad \frac{\varphi \in \mathrm{PA}}{\mathrm{PA} \vdash_{\mathcal{H}} \varphi}$$

Elements from Rautenberg, Troelstra and Schwichtenberg, as well as both.

Definition (Extended Signature of Peano Arithmetic (EPA), simplified)

In addition to the symbols of PA, EPA contains the following function symbols:

$[\,]$ (nil)	$\lvert\ell\rvert$ (length)	$\ell +\!\!+ \ell'$ (append)
$x :: \ell$ (cons)	$\ell[i]$ (indexed access)	$x \rightsquigarrow y$ (implication)

Further, EPA adds the unary predicate symbol $\mathcal{A}$ to PA.

- $\text{EPA} \vdash \overline{\varphi \to \psi} = \overline{\varphi} \rightsquigarrow \overline{\psi}$ (object level implication function)
- If $\varphi \in \mathcal{H}$, then $\text{EPA} \vdash \mathcal{A}\,(\forall x_1. \dots x_n. \varphi)$
- If $\varphi \in \text{PA}$, then $\text{EPA} \vdash \mathcal{A}\,\varphi$

Formal proofs: Spelling out (some of) the Details

Definition (Formal proofs)

A proof of φ is a nonempty list $\ell = [\psi_1, \ldots, \psi_n] : \mathcal{L}(\mathbb{F})$ with $\varphi = \psi_n$ s.t. for each i

- ψ_i is an axiom of PA, a generalisation of a Hilbert axiom, or
- there are $j, j' < i$ such that ψ_i follows from $\psi_j, \psi_{j'}$ by modus ponens.

Definition (Provability predicate)

$$\begin{aligned} \mathrm{Prf}(x, y) &:= (\exists z.\, |x| = S\, z \wedge x[z] = y) \wedge \forall i.\, i < |x| \rightarrow \mathrm{WellFormed}(x, i) \\ \mathrm{WellFormed}(x, i) &:= \mathcal{A}(x) \vee \exists j\, j'.\, j < i \wedge j' < i \wedge x[j] = x[j'] \tilde{\rightarrow} x[i] \end{aligned}$$

Technical Background: Gödel Numberings

Problem

Let $\varphi(x), \psi : \mathbb{F}$.
We used $\varphi(\overline{\psi})$ for 'substituting some encoding of ψ for x in φ'.
ψ is not a **number**, but a **formula**.

Typical issue. Gödel faced it himself.

Remark (Gödelisation)

There are functions $\text{göd} : \mathbb{F} \to \mathbb{N}$, $\text{göd}^{-1} : \mathbb{N} \to \mathbb{F}$ *inverting each other.*

$\varphi(\overline{\psi}) \rightsquigarrow \varphi(\overline{\text{göd}(\psi)})$

Technical Background: Mostowski's Modification

There is $\mathrm{Pr}(x)$ such that $\mathrm{PA} \vdash \varphi$ iff $\mathrm{PA} \vdash \mathrm{Pr}(\overline{\varphi})$.

Definition (Mostowski's Modification, 1965 (modified slightly))

$\mathrm{Pr}^M(x) := \mathrm{Pr}(x) \wedge x \neq \overline{\bot}$.

Lemma

$\mathrm{PA} \vdash \varphi$ *iff* $\mathrm{PA} \vdash \mathrm{Pr}^M(\overline{\varphi})$.

Proof.

- Suppose that $\mathrm{PA} \vdash \varphi$
  1. Observe that $\mathrm{PA} \vdash \mathrm{Pr}(\overline{\varphi})$
  2. In the meta-level, we know that PA is consistent, so $\varphi \neq \bot$ (and thus $\overline{\varphi} \neq \overline{\bot}$)
  3. PA decides equalities: $\mathrm{PA} \vdash \overline{\varphi} = \overline{\bot}$ or $\mathrm{PA} \vdash \overline{\varphi} \neq \overline{\bot}$
  4. By soundness and 2., $\mathrm{PA} \nvdash \overline{\varphi} = \overline{\bot}$
  5. From 1.,3. and 4., conclude $\mathrm{PA} \vdash \mathrm{Pr}(\overline{\varphi}) \dot{\wedge} \overline{\varphi} \neq \overline{\bot}$

Technical Background: Mostowski's Modification

There is $\mathrm{Pr}(x)$ such that $\mathrm{PA} \vdash \varphi$ iff $\mathrm{PA} \vdash \mathrm{Pr}(\overline{\varphi})$.

Definition (Mostowski's modification, 1965 (modified slightly))

$\mathrm{Pr}^M(x) := \mathrm{Pr}(x) \wedge x \neq \overline{\bot}$.

Lemma

$\mathrm{PA} \vdash \neg\mathrm{Pr}^M(\overline{\bot})$.

Proof.

Argue inside PA.

- After introducing, have to show $\mathrm{PA}, \mathrm{Pr}(\overline{\bot}), \overline{\bot} \neq \overline{\bot} \vdash \bot$
- By applying the assumption, left to show $\mathrm{PA}, \mathrm{Pr}(\overline{\bot}), \overline{\bot} \neq \overline{\bot} \vdash \overline{\bot} = \overline{\bot}$
- This follows by reflexivity

□

Technical Background: $\mathsf{CT}_{\mathsf{PA}}$ is too Weak

Axiom ($\mathsf{CT}_{\mathsf{PA}}$)

For every $f : \mathbb{N} \to \mathbb{N}$, there is a formula $\varphi(x_1, x_2)$ such that for all $n : \mathbb{N}$

$$\mathsf{PA} \vdash \forall y.\, \varphi(\overline{n}, y) \leftrightarrow y = \overline{f\ n}.$$

Example

Suppose the successor function $\mathtt{S} : \mathbb{N} \to \mathbb{N}$ is represented by $\varphi_{\mathtt{S}}(x, y)$.

Question: Can we derive, for all $n : \mathbb{N}$, that $\mathsf{PA} \vdash \varphi_{\mathtt{S}}(\overline{n}, \mathsf{S}\,\overline{n})$? **Yes!**

- Use property of $\varphi_{\mathtt{S}}$: $\mathsf{PA} \vdash \mathsf{S}\,\overline{n} = \overline{\mathtt{S} n}$
- By definition of numerals, $\mathsf{PA} \vdash \mathsf{S}\,\overline{n} = \mathsf{S}\,\overline{n}$, easy to finish

Question: Can we derive $\mathsf{PA} \vdash \forall x.\, \varphi_{\mathtt{S}}(x, \mathsf{S}\,x)$? **No!**

- Introduce x: $\mathsf{PA} \vdash \varphi_{\mathtt{S}}(x, \mathsf{S}\,x)$. No way to continue as x not a numeral

Technical Background: Diagonal Lemma

- Functions $\text{diag} := \lambda\varphi.\, \varphi(\overline{\varphi})$, and $\text{diag}_{\mathbb{N}} := \lambda n.\, \text{göd}(\text{diag}(\text{göd}^{-1}(n)))$

Proof.

- Suppose $\varphi(x)$. To find: G such that $\text{PA} \vdash G \leftrightarrow \varphi(\overline{G})$
- Plug $\text{diag}_{\mathbb{N}}$ into CT_{PA}, get $\text{dg}(x, y)$ with $\forall n : \mathbb{N}.\, \text{PA} \vdash \forall x.\, \text{dg}(\overline{n}, x) \leftrightarrow x \equiv \overline{\text{diag}_{\mathbb{N}}\ n}$
- Define $G' := \exists y.\, \text{dg}(x, y) \wedge \varphi(y)$ and $G := G'(\overline{G'})$
- Argue inside PA that

$$
\begin{aligned}
G = G'(\overline{G'}) &= \exists y.\, \text{dg}(\overline{G'}, y) \wedge \varphi(y) \\
&\leftrightarrow \exists y.\, y \equiv \overline{\text{diag}_{\mathbb{N}}(\text{göd}(G'))} \wedge \varphi(y) \\
&\leftrightarrow \exists y.\, y \equiv \overline{\text{göd}(G)} \wedge \varphi(y) \\
&\leftrightarrow \varphi(\overline{G})
\end{aligned}
$$

□

Technical background: Tarski's Theorem

Theorem (Tarski's theorem)

There is no $\mathrm{True}(x) : \mathbb{F}$ *such that for all formulae* φ

$$(\mathbb{N} \vDash \varphi \to \mathbb{N} \vDash \mathrm{True}(\overline{\varphi})) \textit{ and } (\mathbb{N} \nvDash \varphi \to \mathbb{N} \vDash \neg\mathrm{True}(\overline{\varphi})).$$

Proof.

- Suppose $\mathrm{True}(x)$ has this property
- By diagonal lemma and soundness, find G such that $\mathbb{N} \vDash G \leftrightarrow \neg\mathrm{True}(\overline{G})$
- Case distinction
  - If $\mathbb{N} \vDash G$, then $\mathbb{N} \vDash \mathrm{True}(\overline{G})$
    Further, $\mathbb{N} \vDash \neg\mathrm{True}(\overline{G})$ from $\mathbb{N} \vDash G \leftrightarrow \neg\mathrm{True}(\overline{G})$, i.e. $\mathbb{N}$ is inconsistent
  - If $\mathbb{N} \nvDash G$, have $\mathbb{N} \vDash \neg\mathrm{True}(\overline{G})$
    Show $\mathbb{N} \vDash G$. Easy from $\mathbb{N} \vDash G \leftrightarrow \neg\mathrm{True}(\overline{G})$ $\square$

Technical Background: Gödel's First Incompleteness Theorem

Theorem (Strong separability, cf. [HK23])

Suppose $P, Q : \mathbb{N} \to \mathbb{P}$ *are*

- *both semi-decidable and*
- *disjoint (i.e. for all* $n : \mathbb{N}$*, we have* $P\ n \to Q\ n \to \bot$*).*

Then, there is a formula $\varphi(x)$ *such that for all* $n : \mathbb{N}$ *we have*

$$(P\ n \to \mathrm{PA} \vdash \varphi(\overline{n})) \text{ and } (Q\ n \to \mathrm{PA} \vdash \neg\varphi(\overline{n})).$$

Corollary

We find $\mathrm{SProv}(x) : \mathbb{F}$ *such that for all formulas* φ

$$(\mathrm{PA} \vdash \varphi \to \mathrm{PA} \vdash \mathrm{SProv}(\overline{\varphi})) \land (\mathrm{PA} \vdash \neg\varphi \to \mathrm{PA} \vdash \neg\mathrm{SProv}(\overline{\varphi}))$$

Technical background: Gödel's First Incompleteness Theorem (Continued)

We have $\mathrm{SProv}(x)$ such that for all formulae φ

$$(\mathrm{PA} \vdash \varphi \rightarrow \mathrm{PA} \vdash \mathrm{SProv}(\overline{\varphi})) \wedge (\mathrm{PA} \vdash \neg\varphi \rightarrow \mathrm{PA} \vdash \neg\mathrm{SProv}(\overline{\varphi}))$$

Proof (of Gödel's first incompleteness theorem).

- Need to find: Sentence G with $\mathrm{PA} \nvdash G$ and $\mathrm{PA} \nvdash \neg G$
- Plug $\neg\mathrm{SProv}(x)$ into diagonal lemma, obtain $\mathrm{PA} \vdash G \leftrightarrow \neg\mathrm{SProv}(\overline{G})$
- If $\mathrm{PA} \vdash G$
  - Obtain $\mathrm{PA} \vdash \mathrm{SProv}(\overline{G})$ by property of $\mathrm{SProv}(x)$
  - Observe that $\mathrm{PA} \vdash \neg\mathrm{SProv}(\overline{G})$ from diagonal lemma, contradiction
- If $\mathrm{PA} \vdash \neg G$
  - Obtain $\mathrm{PA} \vdash \neg\mathrm{SProv}(\overline{G})$ by property of $\mathrm{SProv}(x)$
  - Observe that $\mathrm{PA} \vdash G$ from diagonal lemma, contradiction

□