Löb's Theorem and Provability Predicates in Coq

Second Bachelor Seminar Talk

Janis Bailitis

Universität des Saarlandes
Advisors: Dr. Yannick Forster, Dr. Dominik Kirst
Supervisor: Prof. Dr. Gert Smolka
Programming Systems Lab

June 6, 2024

What is Löb's theorem?

Problem (Henkin, 1952)

- Assume sufficiently strong formal system, e.g. Peano arithmetic
- There is a sentence S expressing own provability

Question: S independent or provable?

- Kreisel 1953: It depends on provability predicate
  - Only inspected restricted set of provability predicates
- Löb 1955: S is provable
  - But for a strong notion of provability predicate

Provability

Theorem (Löb's theorem, 1955)

Let $\mathrm{Pr}(x)$ *express provability internally in* T*. For all sentences* φ*, we have*

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

$\mathrm{Pr}(x)$ expresses provability **externally** in T if at least $T \vdash \varphi$ implies $T \vdash \mathrm{Pr}(\overline{\varphi})$

- Internal provability predicates require much technical detail
- Can we do this more abstractly in the spirit of Kirst & Peters (2023)?
  - Löb isolated abstract axioms
  - Crucial ones became known as Hilbert-Bernays-Löb derivability conditions
  - Hide most technical details

The Hilbert-Bernays-Löb derivability conditions

Hilbert-Bernays-Löb (HBL) derivability conditions (see Löb, 1955)

We say that $\mathrm{Pr}(x)$ satisfies

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

- Key behind Gödel's second incompleteness theorem
- Many similar conditions analysed by Kurahashi (2020, 2021)
- Also used in mechanisations with proof assistants (Isabelle)
  - ➤ Paulson (2015) proves first two; uses hereditarily finite set theory
  - ➤ Assumed without proof by Popescu & Traytel (2021)

The Hilbert-Bernays-Löb derivability conditions

We use Peano arithmetic (PA).

Hilbert-Bernays-Löb (HBL) derivability conditions (see Löb, 1955)

We say that $\mathrm{Pr}(x)$ satisfies

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

- We analysed how far Church's Thesis for arithmetic ($\mathrm{CT}_{\mathrm{PA}}$)[1] takes us
  - ➤ Gives more abstract formula
  - ➤ Investigate which axioms hold

[1]Formalised for first-order arithmetic by Hermes and Kirst 2022, proven consistent by Kirst and Peters 2023.

Defining a first provability candidate

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

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

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

Lemma (Weak representability, cf. Hermes & Kirst (2023))

Let $P : \mathbb{N} \to \mathbb{P}$ be enumerable. There is a formula $\varphi(x)$ such that for all $n : \mathbb{N}$

$$P\ n \textit{ iff } \mathsf{PA} \vdash \varphi(\overline{n}).$$

We inspect $\lambda\varphi.\, \mathsf{PA} \vdash \varphi$, enumerability mechanised by Forster, Kirst & Smolka (2019).

Corollary

We obtain $\mathrm{Pr}(x)$ such that $\mathsf{PA} \vdash \varphi$ iff $\mathsf{PA} \vdash \mathrm{Pr}(\overline{\varphi})$.

$\mathrm{Pr}(x)$ is **external** provability predicate.

A problem arises

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

Theorem (Gödel's second incompleteness theorem, 1931)

Suppose $\mathrm{Pr}(x)$ *is an internal provability predicate for* PA. *Then,* $\mathrm{PA} \nvdash \dot{\neg}\mathrm{Pr}(\overline{\dot{\bot}})$.

HBL derivability conditions imply this theorem via Löb's Theorem.

Definition (Mostowski's modification, 1965 (modified slightly[1]))

We set $\mathrm{Pr}'(x) := \mathrm{Pr}(x) \mathbin{\dot{\wedge}} x \neq \overline{\dot{\bot}}$.

Lemma

$\mathrm{PA} \vdash \varphi$ *iff* $\mathrm{PA} \vdash \mathrm{Pr}'(\overline{\varphi})$, *i.e.* $\mathrm{Pr}'(x)$ *is an external provability predicate.*

Lemma

$\mathrm{PA} \vdash \dot{\neg}\mathrm{Pr}'(\overline{\dot{\bot}})$.

[1] This formulation is mentioned by Bezboruah & Shepherdson (1976).

Formal proofs in Peano Arithmetic: Getting internal

Problem

- $\mathrm{Pr}'(x)$ is external provability predicate
- We showed $\mathrm{PA} \vdash \dot{\neg}\mathrm{Pr}'(\overline{\dot{\bot}})$
- If Pr' satisfied the HBL conditions, we would have $\mathrm{PA} \nvdash \dot{\neg}\mathrm{Pr}'(\overline{\dot{\bot}})$

$\rightarrow$ Not all external provability predicates satisfy the HBL conditions

Towards a solution

PA should reason about proofs: Find formula $\mathrm{Prf}(x, y)$ expressing 'x is a proof of y'.

Remark (Formal proofs, cf. [Rau10])

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 *or*
- *there are $j, j' < i$ such that ψ_i follows from $\psi_j, \psi_{j'}$ by modus ponens.*

Obtaining the required functions inside Peano Arithmetic

If PA has a formula $\mathrm{Prf}(x, y)$ capturing proofs, $\mathrm{Pr}(y) := \dot{\exists} x.\, \mathrm{Prf}(x, y)$ is internal.

- Gödel, Boolos, Smith and Rautenberg define $\mathrm{Prf}(x, y)$ from scratch
  - E.g. they explicitly define list formulas such as $\mathrm{len}(x, y)$ for length inside PA

Our approach: Use $\mathrm{CT}_{\mathrm{PA}}$ to obtain these formulas abstractly

Next problem (terms and numerals), simplified

1. We know: For all $\ell : \mathcal{L}(\mathbb{F})$ that $\mathrm{PA} \vdash \dot{\forall} z.\, \mathrm{len}(\bar{\ell}, z) \dot{\leftrightarrow} z = \overline{|\ell|}$
2. Suppose that $\mathrm{Prf}(x, y)$ uses $\mathrm{len}(x, z)$ internally for some z
3. When we destruct $\dot{\exists} x.\, \mathrm{Prf}(x, y)$, we only know $\mathrm{len}(x, z)$ for some **terms** x, z
4. x not necessarily a **numeral**: We cannot apply property from 1.

Introducing the required functions through axioms

New idea needed: Add functions and constants to PA.

$[]$ (nil)	$\lvert\ell\rvert$ (length)	$\ell +\!\!+ \ell'$ (append)
$x :: \ell$ (cons)	$\ell[i]$ (indexed access)	$x \tilde{\to} y$ (implication)

These functions should have standard properties, we assume them through axioms.

Required properties, not exhaustive

- $\mathsf{PA} \vdash \lvert[]\rvert = \mathsf{O}$
- $\mathsf{PA} \vdash \dot\forall \ell\, x.\ \lvert x :: \ell\rvert = \mathsf{S}\lvert\ell\rvert$
- $\mathsf{PA} \vdash \dot\forall \ell\, x.\ (x :: \ell)[\mathsf{O}] = x$
- $\mathsf{PA} \vdash \dot\forall i\, \ell\, x.\ (x :: \ell)[\mathsf{S}i] = \ell[i]$
- $\mathsf{PA} \vdash \dot\forall \ell\, \ell'.\ \lvert \ell +\!\!+ \ell' \rvert = \lvert\ell\rvert + \lvert\ell'\rvert$
- $\mathsf{PA} \vdash \dot\forall i\, \ell\, \ell'.\ i < \lvert\ell\rvert \dot\to \ell[i] = (\ell +\!\!+ \ell')[i]$
- $\mathsf{PA} \vdash \dot\forall i\, \ell\, \ell'.\ i < \lvert\ell'\rvert \dot\to \ell'[i] = (\ell +\!\!+ \ell')[i + \lvert\ell\rvert]$
- For all formulas φ, ψ, we have $\mathsf{PA} \vdash \overline{\varphi \dot\to \psi} = \overline{\varphi} \tilde{\to} \overline{\psi}$

The Hilbert-Bernays-Löb derivability conditions

Hilbert-Bernays-Löb derivability conditions

We say that $\mathrm{Pr}(x)$ satisfies

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

Reminder (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 or
- there are $j, j' < i$ such that ψ_i follows from $\psi_j, \psi_{j'}$ by modus ponens.

Proof (of the distributivity law).

Let ℓ, ℓ' be proofs of $\varphi \dot{\to} \psi$ and φ, respectively. Then, $(\ell + \ell') + [\psi]$ proves ψ. □

The Hilbert-Bernays-Löb derivability conditions

Hilbert-Bernays-Löb derivability conditions

We say that $\mathrm{Pr}(x)$ satisfies

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

Reminder (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 or
- there are $j, j' < i$ such that ψ_i follows from $\psi_j, \psi_{j'}$ by modus ponens.

- Necessitation not obvious: How to formalise ND-derivations this way?
- Need equivalent system with only modus ponens and axioms $\rightsquigarrow$ Hilbert system

Deriving internal necessitation

Definition (Σ_1-formulas, Kleene (1943) & Mostowski (1947))

Let φ and $\psi(x_1, \ldots, x_n)$ be formulas such that

- $\varphi = \dot{\exists} x_1 \ldots \dot{\exists} x_n .\, \psi(x_1, \ldots, x_n)$ and
- ψ does not use quantifiers (except bounded ones, e.g. $\dot{\forall} y < x$)

We say that φ is a **Σ_1-formula**.

We can show that our internal provability predicate is Σ_1.

Theorem (Provable Σ_1-completeness, cf. [Rau10])

Let φ be a Σ_1-formula. Then $\mathrm{PA} \vdash \varphi \dot{\to} \mathrm{Pr}(\overline{\varphi})$*.*

- In the proof, provability on open formulas is needed (e.g. $\mathrm{Pr}(\overline{\varphi(x_1, \ldots, x_n)})$)
- For instance, we want to obtain $\mathrm{Pr}(\overline{\varphi(\overline{42}, \ldots, \overline{42})})$ from $\mathrm{Pr}(\overline{\varphi(x_1, \ldots, x_n)})$ by substitution

Achievements and further work

We have

- Mechanised the abstract proof of Löb's Theorem from the HBL conditions
- Mechanised the Diagonal Lemma as well as important corollaries such as Gödel's first incompleteness theorem
- Analysed why external provability predicates are too weak for Löb's Theorem
- Developed a rough understanding of internal provability predicates

We will

- Formalise the needed Hilbert system for the necessitation proof
- Understand internal necessitation sufficiently
- Finish mechanising the results in Coq
- Optional: Resolve identified weaknesses of Coq library for first-order logic [Kir+22]

Thanks for your attention.

[BS76] A. Bezboruah and J. C. Shepherdson. **'Gödel's Second Incompleteness Theorem for Q'.** In: **The Journal of Symbolic Logic** 41.2 (1976), pp. 503–512. ISSN: 00224812. URL: `http://www.jstor.org/stable/2272251`.

[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`.

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

[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.

[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.

[KH23] Dominik Kirst and Marc Hermes. **'Synthetic Undecidability and Incompleteness of First-Order Axiom Systems in Coq'.** In: **J. Autom. Reason.** 67.1 (2023), p. 13. DOI: 10.1007/S10817-022-09647-X.

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

[Kle43] S. C. Kleene. **'Recursive Predicates and Quantifiers'.** In: **Transactions of the American Mathematical Society** 53 (1943), pp. 41–73. DOI: 10.2307/2267986.

[KP23] Dominik Kirst and Benjamin Peters. **'Gödel's Theorem Without Tears - Essential Incompleteness in Synthetic Computability'.** 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.

[Kur20] Taishi Kurahashi. **'A note on derivability conditions'.** In: **The Journal of Symbolic Logic** 85.3 (Sept. 2020), 1224–1253. ISSN: 1943-5886. DOI: 10.1017/jsl.2020.33.

[Kur21] Taishi Kurahashi. **'Rosser Provability and the Second Incompleteness Theorem'.** In: **Advances in Mathematical Logic**. Ed. by Toshiyasu Arai et al. Singapore: Springer Nature Singapore, 2021, pp. 77–97. ISBN: 978-981-16-4173-2.

[Lö55] M. H. Löb. **'Solution of a Problem of Leon Henkin'.** In: **The Journal of Symbolic Logic** 20.2 (1955), pp. 115–118. ISSN: 00224812. URL: http://www.jstor.org/stable/2266895.

[Mos47] Andrzej Mostowski. **'On definable sets of positive integers'.** In: **Fundamenta Mathematicae** 34.1 (1947), pp. 81–112. URL: http://eudml.org/doc/213118.

[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.

[Pau15] Lawrence C. Paulson. **'A Mechanised Proof of Gödel's Incompleteness Theorems Using Nominal Isabelle'.** In: **Journal of Automated Reasoning** 55 (2015), pp. 1–37. DOI: 10.1007/s10817-015-9322-8.

[PT21] A. Popescu and D. Traytel. **'Distilling the Requirements of Gödel's Incompleteness Theorems with a Proof Assistant'.** In: **Journal of Automated Reasoning** (2021). DOI: 10.1007/s10817-021-09599-8.

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

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** $\to$ $\overline{\psi}$ not a numeral.

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.*

$\varphi(\overline{\psi}) \rightsquigarrow \varphi(\overline{\text{göd}(\psi)})$
We also assume a Gödel numbering $\text{göd}_{\mathcal{L}}$ for lists of formulas; $\varphi(\bar{\ell}) \rightsquigarrow \varphi(\overline{\text{göd}_{\mathcal{L}}(\ell)})$

Technical background: External provability is too weak

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

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

We set $\mathrm{Pr}'(x) := \mathrm{Pr}(x) \mathbin{\dot\wedge} x \neq \overline{\dot\bot}$.

Lemma

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

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 $\varphi \neq \dot\bot$ (and thus $\overline{\varphi} \neq \overline{\dot\bot}$)
  3. PA decides equalities: $\mathrm{PA} \vdash \overline{\varphi} = \overline{\dot\bot}$ or $\mathrm{PA} \vdash \overline{\varphi} \neq \overline{\dot\bot}$
  4. By soundness and 2., $\mathrm{PA} \nvdash \overline{\varphi} = \overline{\dot\bot}$
  5. From 1.,3. and 4., conclude $\mathrm{PA} \vdash \mathrm{Pr}(\overline{\varphi}) \mathbin{\dot\wedge} \overline{\varphi} \neq \overline{\dot\bot}$

Technical background: External provability is too weak

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

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

We set $\mathrm{Pr}'(x) := \mathrm{Pr}(x) \mathbin{\dot\wedge} x \neq \overline{\dot\bot}$.

Lemma

$\mathrm{PA} \vdash \dot\neg \mathrm{Pr}'(\overline{\dot\bot})$.

Proof.

We argue inside PA.

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

□

Technical background: External provability is too weak

Lemma (Weak representability, cf. Hermes & Kirst (2023))

Let $P : \mathbb{N} \to \mathbb{P}$ be enumerable. There is a formula $\varphi(x)$ such that for all $n : \mathbb{N}$

$$P\ n \text{ iff } \mathrm{PA} \vdash \varphi(\overline{n}).$$

Recap

- We proved $\mathrm{Pr}'(x)$ weakly represents $\lambda\varphi.\ \mathrm{PA} \vdash \varphi$
- We showed $\mathrm{PA} \vdash \dot{\neg}\mathrm{Pr}'(\overline{\dot{\bot}})$

We can show: If $\mathrm{Pr}'(x)$ was internal, then $\mathrm{PA} \nvdash \dot{\neg}\mathrm{Pr}'(\overline{\dot{\bot}})$, for instance via Gödel's second incompleteness theorem.

Takeaway

Weak representability does not suffice to define internal provability predicates.

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* $\varphi(x, y)$ *such that for all* $n : \mathbb{N}$

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

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 $\mathsf{PA} \vdash \dot\forall x.\, \varphi_{\mathtt{S}}(x, \mathsf{S}\,x)$? **No!**

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

Technical background: $\mathrm{CT}_{\mathrm{PA}}$ is too weak

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

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

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

- Suppose we define $\mathrm{Prf}(x, y)$ using formulas resulting from $\mathrm{CT}_{\mathrm{PA}}$, s.t. e.g. $\varphi(x, z)$ appears in $\mathrm{Prf}(x, y)$
- We have to show $\mathrm{PA} \vdash (\dot{\exists} x.\, \mathrm{Prf}(x, \overline{\varphi \dot{\to} \psi})) \dot{\to} (\dot{\exists} x'.\, \mathrm{Prf}(x', \overline{\varphi})) \dot{\to} \dot{\exists} x''.\, \mathrm{Prf}(x'', \overline{\psi})$
- After introducing, we know $\varphi(x, z)$ and $\varphi(x', z')$ for some terms z, z'
- x, x' not necessarily numerals: Defining property from $\mathrm{CT}_{\mathrm{PA}}$ not applicable

Formal proofs: Spelling out (some of) the details

Reminder (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 or
- there are $j, j' < i$ such that ψ_i follows from $\psi_j, \psi_{j'}$ by modus ponens.

Definition (Provability predicate)

Let $\mathsf{Ax}(x)$ be a formula capturing axioms of PA (details to be investigated).

$$\mathsf{Prf}(x, y) := (\dot{\exists} z.\ |x| = Sz \mathbin{\dot{\wedge}} x[z] = y) \mathbin{\dot{\wedge}} \dot{\forall} i.\ i < |x| \mathbin{\dot{\to}} \mathsf{WellFormed}(x, i)$$

$$\mathsf{WellFormed}(x, i) := \mathsf{Ax}(x) \mathbin{\dot{\vee}} \dot{\exists} j\, j'.\ j < i \mathbin{\dot{\wedge}} j' < i \mathbin{\dot{\wedge}} x[j] = x[j'] \mathbin{\tilde{\to}} x[i]$$

Technical background: Provability on open formulas

Problem

- Let φ be a formula with free variables $x_1, \ldots, x_n$, i.e. $\varphi = \varphi(x_1, \ldots, x_n)$
- $\Pr(\overline{\varphi(x_1, \ldots, x_n)})$ is closed. $x_1, \ldots, x_n$ are hard-coded in encoding
- We want to obtain $\Pr(\overline{\varphi(\overline{42}, \ldots, \overline{42})})$ from $\Pr(\overline{\varphi(x_1, \ldots, x_n)})$ by substitution

First attempt to circumvent this issue, cf. [Rau10]

- It is possible to define an n-place function sb_n in PA satisfying
  $$\mathrm{PA} \vdash \mathrm{sb}_n(\overline{\varphi(x_1, \ldots, x_n)}, t_1, \ldots, t_n) = \overline{\varphi(t_1, \ldots, t_n)}$$
  for any n-bounded formula φ and terms $t_1, \ldots, t_n$.
- Then, $\Pr_n(x, y_1, \ldots, y_n) := \Pr(\mathrm{sb}_n(x, y_1, \ldots, y_n))$ has desired property
- Not satisfactory due to dependency on n

Paulson formalised yet another approach in the Isabelle proof assistant.