A Formally Verified Abstract Account of Gödel's Incompleteness Theorems

Andrei Popescu

Middlesex University London

Dmitriy Traytel

ETH zürich

Gödel’s Incompleteness Theorems 1931

Gödel's Incompleteness Theorems 1931

Fix a consistent logical theory that

- contains enough arithmetic,
- can itself be arithmetized.

Gödel's Incompleteness Theorems 1931

Fix a consistent logical theory that

- contains enough arithmetic,
- can itself be arithmetized.

There are sentences that the theory cannot decide (i.e., neither prove nor disprove).

Gödel's Incompleteness Theorems 1931

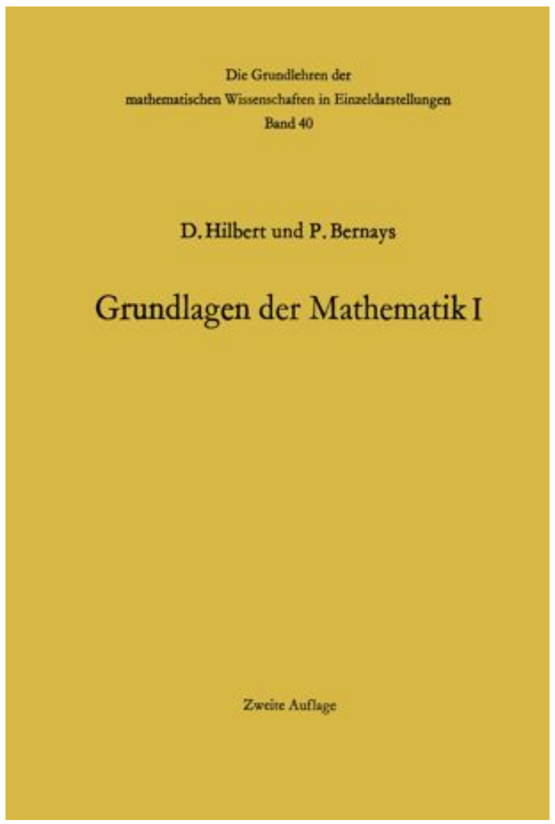
Fix a consistent logical theory that
- contains enough arithmetic,
- can itself be arithmetized.

There are sentences that the theory cannot decide (i.e., neither prove nor disprove).

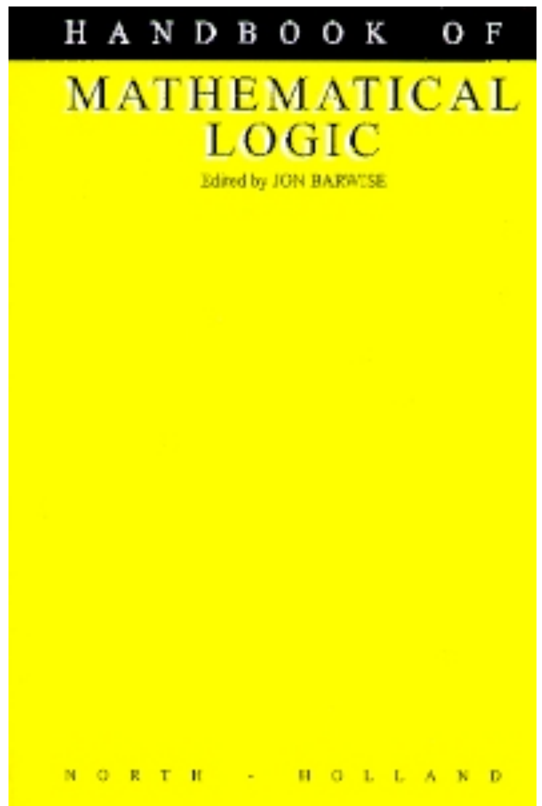
The theory cannot prove (an internal formulation of) its own consistency.

Pen and Paper Proofs of  and 

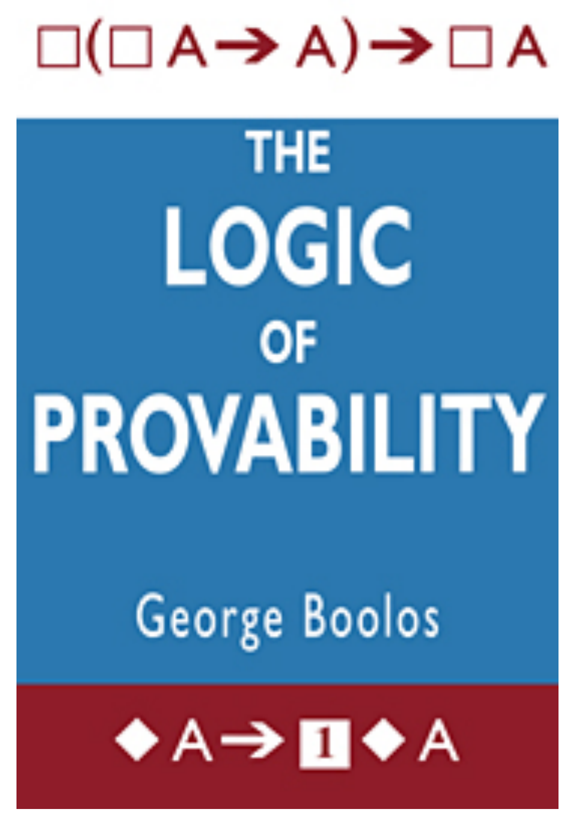
Pen and Paper Proofs of 1 and 2



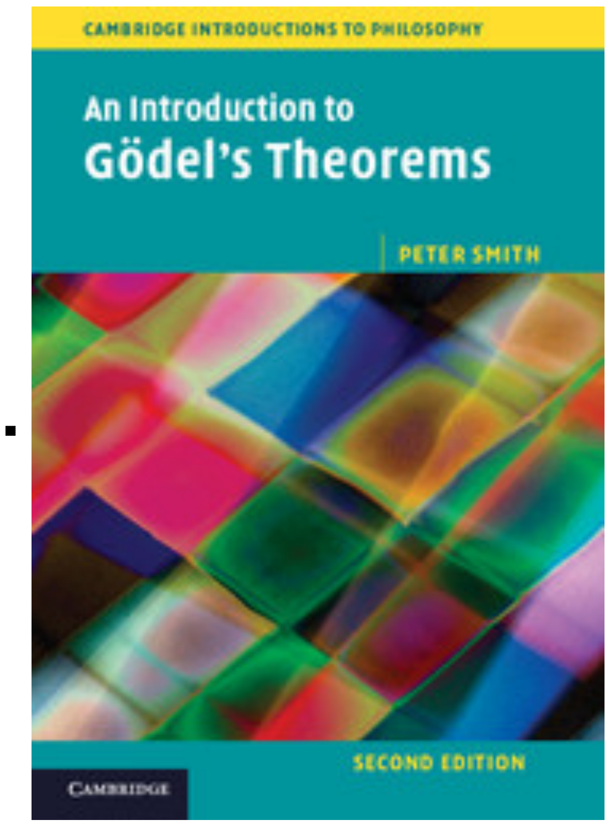
…



…

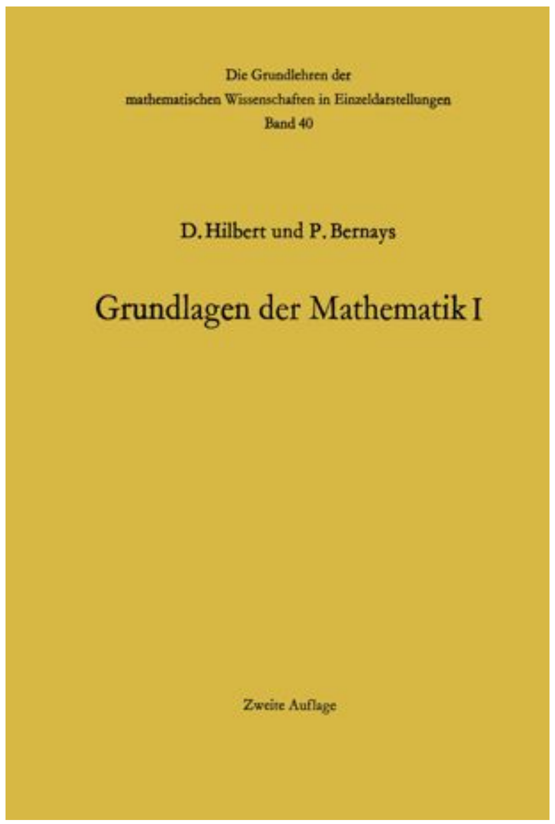


…

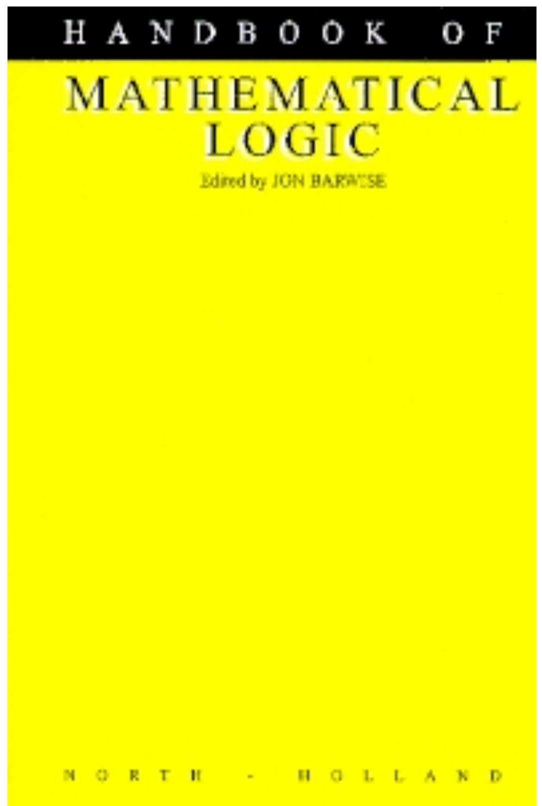


…

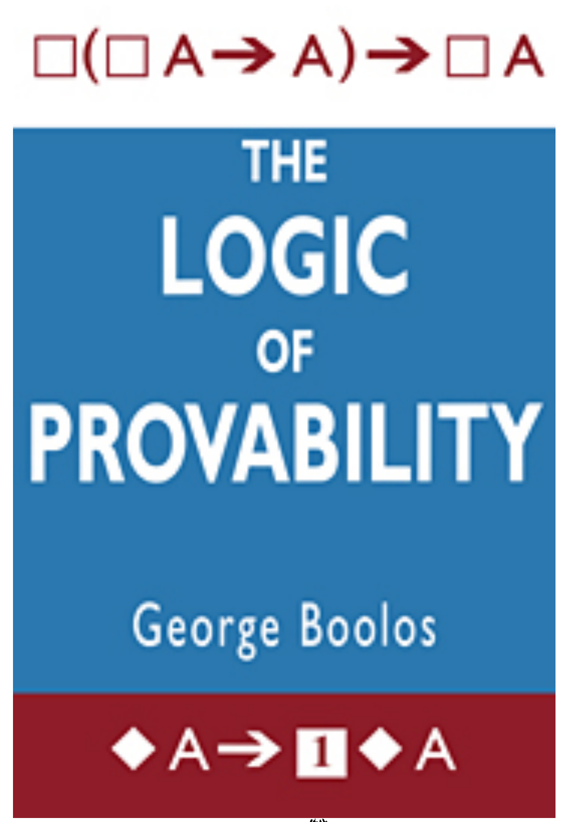
Pen and Paper Proofs of  and 



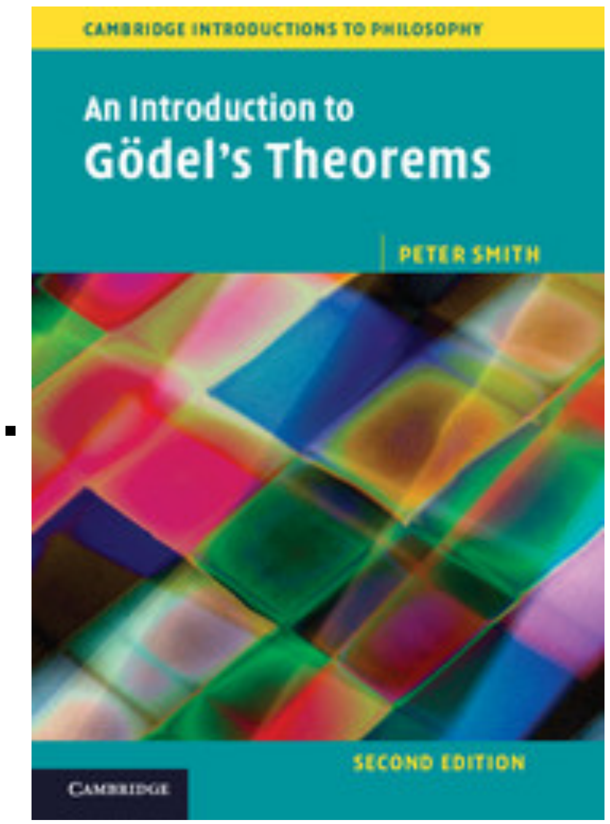
…



…



…

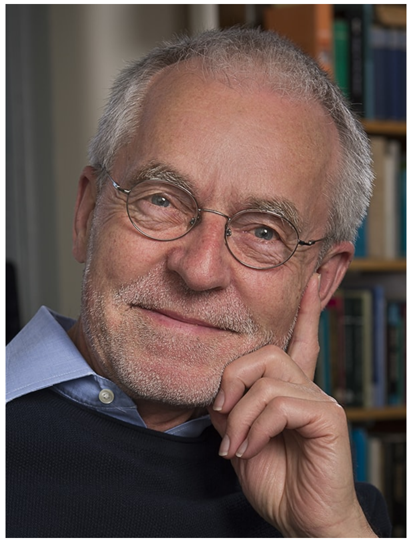


…

The reader who does not like **incomplete and (apparently) irremediably messy proofs** of syntactic facts may wish to skim over the rest of this chapter and take it for granted that …

Formal Verifications of  and 

Formal Verifications of 1 and 2

TEM	NQTHM	HOL Light	Coq	Isabelle
Sieg	Shankar	Harrison	O’Connor	Paulson
				
1978	1986	2004	2005	2015

End of story

End of story?

~~End of story?~~

Formal Verifications of 1 and 2

Shared structure

Formal Verifications of  and 

Shared structure

- Fix a particular logic: **Classical FOL**

Formal Verifications of  and 

Shared structure

- Fix a particular logic: **Classical FOL**
- Fix a particular theory (+ finite extensions of it)
  - **Arithmetic** (Harrison, O’Connor)
  - **Hereditarily finite set theory** (Sieg, Shankar, Paulson)

Formal Verifications of  and 

Shared structure

- Fix a particular logic: **Classical FOL**
- Fix a particular theory (+ finite extensions of it)
    - **Arithmetic** (Harrison, O'Connor)
    - **Hereditarily finite set theory** (Sieg, Shankar, Paulson)

Formal Verifications of  and 

Shared structure

- Fix a particular logic: **Classical FOL**
- Fix a particular theory (+ finite extensions of it)
    - **Arithmetic** (Harrison, O’Connor)
    - **Hereditarily finite set theory** (Sieg, Shankar, Paulson)
- Tour de force for the particular combination

Formal Verifications of 1 and 2

Shared structure

- Fix a particular logic: **Classical FOL**
- Fix a particular theory (+ finite extensions of it)
  - **Arithmetic** (Harrison, O'Connor)
  - **Hereditarily finite set theory** (Sieg, Shankar, Paulson)
- Tour de force for the particular combination

Scope of 1 and 2 remains largely unexploded

Formal Verifications of 1 and 2

Shared structure

- Fix a particular logic: **Classical FOL**
- Fix a particular theory (+ finite extensions of it)
  - **Arithmetic** (Harrison, O'Connor)
  - **Hereditarily finite set theory** (Sieg, Shankar, Paulson)
- Tour de force for the particular combination

Scope of 1 and 2 remains largely unexploded

E.g. do they hold for **Intuitionistic FOL**, **HOL**, **CIC**?

Our Motto:

Our Motto:
Don’t Fix, Gather!

Our Contributions

Our Contributions

- Abstract formalization of  and 
  - **Answer** “What must/may a logic/theory offer?”
  - **Understand** variants and **distill** trade-offs from the literature
  - **Correct** a mistake in a pen and paper proof

Our Contributions

- Abstract formalization of 1 and 2
  - **Answer** “What must/may a logic/theory offer?”
  - **Understand** variants and **distill** trade-offs from the literature
  - **Correct** a mistake in a pen and paper proof

- Concrete instantiation to hereditarily finite set theory
  - **Reproduce** (for 1) and **improve** (for 2) Paulson’s formalization

What must a logic/theory offer?

Generic Syntax	Connectives	Provability Relation	Numerals

What must a logic/theory offer?

Generic Syntax	Connectives	Provability Relation	Numerals

What may a logic/theory offer?

Classical Logic	Order-like Relation	Proofs	Encodings
Represent-ability	Derivability Conditions	Standard Model	Soundness
Consistency	Omega-Consistency	Completeness of Provability	Proofs vs. Provability

Generic
Syntax

- **sets:** Var, Term, Fmla with Var⊆Term

Generic Syntax

- **sets:** Var, Term, Fmla with Var⊆Term
- **operators:**

FV_Term : Term → 2^{Var}

FV : Fmla → 2^{Var}

subst_Term : Term → Var → Term → Term

subst : Fmla → Var → Term → Fmla

Generic Syntax

Generic Syntax

- **sets:** Var, Term, Fmla with Var$\subseteq$Term
- **operators:**

  FV_Term : Term → 2^{Var}

  FV : Fmla → 2^{Var}

  subst_Term : Term → Var → Term → Term

  subst : Fmla → Var → Term → Fmla

- **properties, e.g.:**

  $x \in FV(\phi)$ implies $FV(\text{subst}\ \phi\ x\ s) = FV(\phi) - \{x\} \cup FV\_Term(s)$

Generic Syntax

- **sets:** Var, Term, Fmla with Var$\subseteq$Term
- **operators:**

  FV_Term : Term $\rightarrow$ 2^{Var}

  FV : Fmla $\rightarrow$ 2^{Var}

  subst_Term : Term $\rightarrow$ Var $\rightarrow$ Term $\rightarrow$ Term

  subst : Fmla $\rightarrow$ Var $\rightarrow$ Term $\rightarrow$ Fmla
- **properties, e.g.:**

  $x \in FV(\phi)$ implies $FV(\text{subst } \phi \ x \ s) = FV(\phi) - \{x\} \cup FV\_Term(s)$

We require unary substitution only.
We derive parallel substitution from it.

Connectives

- **operators:**

  $\equiv$: Term → Term → Fmla

  $\to$, $\wedge$, $\vee$: Fmla → Fmla → Fmla

  $\neg$: Fmla → Fmla

  $\bot$, $\top$: Fmla

  $\exists$, $\forall$: Var → Fmla → Fmla

Connectives

Connectives

- **operators:**

  $\equiv$: Term → Term → Fmla

  →, ∧, ∨ : Fmla → Fmla → Fmla

  ¬ : Fmla → Fmla

  ⊥, ⊤ : Fmla

  ∃, ∀ : Var → Fmla → Fmla

We require a minimal list w.r.t. intuitionistic deduction and define the rest.
Note: operators, not constructors

Provability Relation

- **unary relation:**

  $\vdash \subseteq \mathrm{Fmla}$

  we write $\vdash\phi$ if $\phi \in \vdash$

- **properties:**

  $\vdash$ contains the standard (Hilbert-style) intuitionistic FOL axioms about the connectives

- **unary relation:**

  $\vdash \subseteq \text{Fmla}$

  we write $\vdash\phi$ if $\phi \in \vdash$

- **properties:**

  $\vdash$ contains the standard (Hilbert-style) intuitionistic FOL axioms about the connectives

- **nonempty set:**

  $\text{Num} \subseteq \text{Fmla}_0$

Provability Relation

Numerals

- **property:** $\vdash \neg\neg\phi \to \phi$

Classical Logic

- **property:** $\vdash \neg\neg\phi \rightarrow \phi$

- **formula:** $< \in Fmla_2$
- **properties, e.g.:**

for all $\phi \in Fmla_1$ and $n \in Num$,

if $\vdash \phi(m)$ for all $m \in Num$, then $\vdash \forall x.\ x<n \rightarrow \phi(x)$

Classical Logic

Order-like Relation

- **property:** $\vdash \neg\neg\phi \rightarrow \phi$

Classical Logic

- **formula:** $< \in Fmla_2$
- **properties, e.g.:**

for all $\phi \in Fmla_1$ and $n \in Num$,

if $\vdash \phi(m)$ for all $m \in Num$, then $\vdash \forall x.\ x<n \rightarrow \phi(x)$

Order-like Relation

- **set:** Proof
- **binary relation:** $\Vdash\ \in Proof \times Fmla$

we write $p \Vdash \phi$ if $(p,\phi) \in \Vdash$

Proofs

- **operators:**

  $\langle \_ \rangle$: Fmla → Num and $\langle \_ \rangle$: Proof → Num

- **formulas** <u>subst</u>, <u>⊩</u>, <u>¬</u>
- **property:**

  behave like operators/relations (subst, ⊩, ¬) on encodings

Encodings

Represent-
ability

- **operators:**

  $\langle \_ \rangle : \mathrm{Fmla} \to \mathrm{Num}$ and $\langle \_ \rangle : \mathrm{Proof} \to \mathrm{Num}$

Encodings

- **formulas** $\underline{\text{subst}}$, $\underline{\Vdash}$, $\underline{\neg}$
- **property:**

  behave like operators/relations (subst, $\Vdash$, $\neg$) on encodings

Represent-
ability

- **property:** $\nVdash \bot$

Consistency

- **operators:**

  $\langle \_ \rangle : \mathrm{Fmla} \to \mathrm{Num}$ and $\langle \_ \rangle : \mathrm{Proof} \to \mathrm{Num}$

Encodings

- **formulas** $\underline{\mathrm{subst}}$, $\underline{\Vdash}$, $\underline{\neg}$
- **property:**

  behave like operators/relations (subst, $\Vdash$, $\neg$) on encodings

Represent-ability

- **property:** $\nvdash \perp$

Consistency

- **property:** For all $\phi \in \mathrm{Fmla}_1$,

  if $\vdash \neg\phi(n)$ for all $n \in \mathrm{Num}$ then $\nvdash \neg\neg(\exists x.\phi(x))$

Omega-Consistency

What must a logic/theory offer?

Generic Syntax	Connectives	Provability Relation	Numerals

What may a logic/theory offer?

Classical Logic	Order-like Relation	Proofs	Encodings
Represent-ability	Derivability Conditions	Standard Model	Soundness
Consistency	Omega-Consistency	Completeness of Provability	Proofs vs. Provability

Omega-Consistency

Proofs

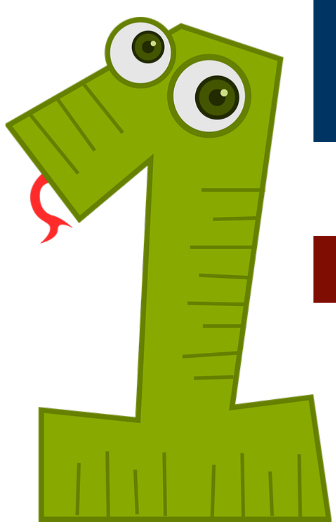
<u>subst</u>, $\underline{\Vdash}$

$\vdash \phi$ implies $\vdash \underline{\vdash} \langle \phi \rangle$

Represent-ability

Derivability Conditions

Encodings



There exists $\phi \in \mathrm{Fmla}_0$ such that $\nvdash \phi$ and $\nvdash \neg \phi$

Rosser's Trick

Consistency

Proofs

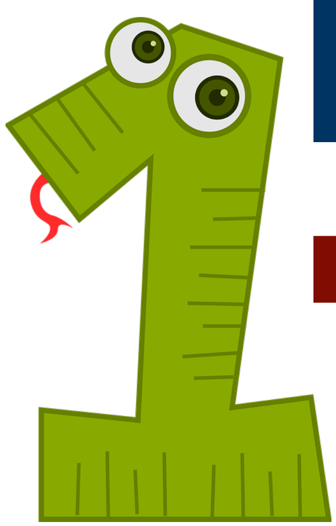
<u>subst, $\Vdash$</u>

Represent-ability

$\vdash \phi$ implies $\vdash \underline{\vdash} \langle \phi \rangle$

Derivability Conditions

Encodings



a la Rosser

There exists $\phi \in Fmla_0$ such that

$\nvdash \phi$ and $\nvdash \neg\phi$

Rosser's Trick

Consistency

Proofs

$\underline{\text{subst}}$, $\underline{\Vdash}$

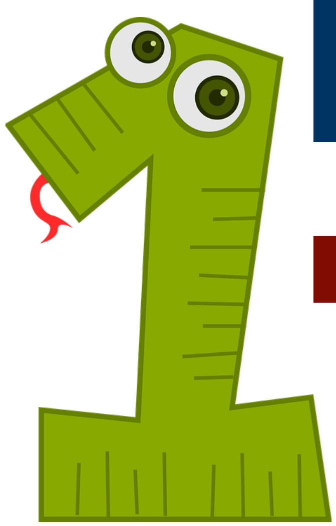
$\vdash\phi$ implies $\vdash\underline{\vdash}\langle\phi\rangle$

Represent-ability

$\underline{\neg}$

Derivability Conditions

Encodings



a la Rosser

There exists $\phi \in \mathrm{Fmla}_0$ such that

$\nvdash \phi$ and $\nvdash \neg\phi$

Rosser's Trick

Consistency

Proofs

Order-like Relation

$\underline{\text{subst}}$, $\underline{\Vdash}$

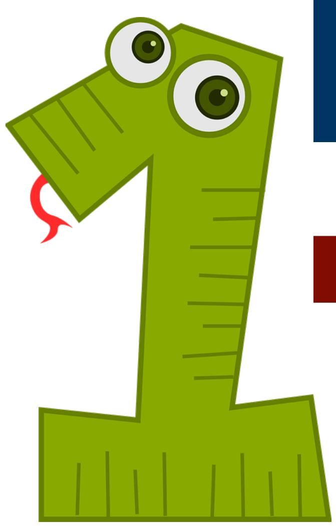
Represent-ability

$\underline{\neg}$

$\vdash \phi$ implies $\vdash \underline{\vdash} \langle \phi \rangle$

Derivability Conditions

Encodings



a la Rosser

There exists $\phi \in \mathrm{Fmla}_0$ such that $\nvdash \phi$ and $\nvdash \neg\phi$

Standard Model

Soundness

Completeness of Provability

Proofs vs. Provability

<u>subst</u>

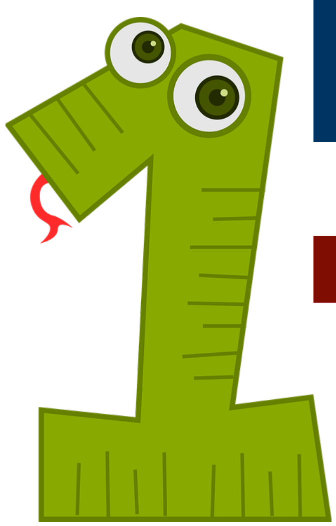
Represent-ability

$\vdash \phi$ implies $\vdash \underline{\vdash}\langle \phi \rangle$

Derivability Conditions

$\vdash \underline{\vdash}\langle \phi \rangle$ implies $\vdash \phi$

Encodings



semantic

There exists $\phi \in \mathrm{Fmla}_0$ such that

$\nvdash \phi$ and $\nvdash \neg\phi$

and ϕ is true in the standard model

Consistency

Classical Logic

<u>subst</u>

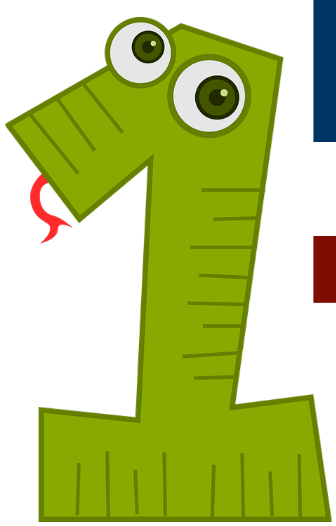
Representability

$\vdash \phi$ implies $\vdash \underline{\vdash}\langle \phi \rangle$

Derivability Conditions

$\vdash \underline{\vdash}\langle \phi \rangle$ implies $\vdash \phi$

Encodings



classical

There exists $\phi \in \mathrm{Fmla}_0$ such that $\nvdash \phi$ and $\nvdash \neg \phi$

Consistency

$\vdash \underline{\vdash}\langle\phi\rangle \to \underline{\vdash}\langle\underline{\vdash}\langle\phi\rangle\rangle$

$\vdash \underline{\vdash}\langle\phi\rangle \wedge \underline{\vdash}\langle\phi \to \psi\rangle \to \underline{\vdash}\langle\psi\rangle$

$\vdash\phi$ implies $\vdash \underline{\vdash}\langle\phi\rangle$

<u>subst</u>

Represent-
ability

Derivability
Conditions

Encodings

$$\nvdash \neg\underline{\vdash}\langle\bot\rangle$$

In the paper:
Jeroslow's
“improvement"
to remove this
condition

results in weaker
conclusion
\+ mistake in proof

Consistency

$\vdash \underline{\vdash}\langle\phi\rangle \to \underline{\vdash}\langle\underline{\vdash}\langle\phi\rangle\rangle$

$\vdash \underline{\vdash}\langle\phi\rangle \wedge \underline{\vdash}\langle\phi \to \psi\rangle \to \underline{\vdash}\langle\psi\rangle$

$\vdash\phi$ implies $\vdash \underline{\vdash}\langle\phi\rangle$

<u>subst</u>

Represent-
ability

Derivability
Conditions

Encodings

$$\nvdash \neg \underline{\vdash}\langle\bot\rangle$$

Summary Using our generic infrastructure (Section 2), we have formally proved several abstract incompleteness results. They include four versions of $\mathcal{IT}_1$:

- Gödel's original $\mathcal{IT}_1$ (Theorem 9) and an $\mathcal{IT}_1$ based on classical logic (Theorem 12) required the formalization of some well-known arguments without change.
- Rosser's $\mathcal{IT}_1$ (Theorem 10) involved the generalization of a well-known argument: distilling two abstract conditions, Ord_1 and Ord_2.
- Novel semantic variants of $\mathcal{IT}_1$ (Theorems 11 and 13) were born from abstractly connecting standard models, HBL_1's "iff" version, and proof representability.

They also include two versions of $\mathcal{IT}_2$:

- The standard $\mathcal{IT}_2$ based on the three derivability conditions (Theorem 14) again only required formalizing a well-known argument.
- The alternative, Jeroslow-style $\mathcal{IT}_2$ (Theorems 17 and 18) involved a detailed analysis and correction of an existing abstract result.

Summary Using our generic infrastructure (Section 2), we have formally proved several abstract incompleteness results. They include four versions of $\mathcal{IT}_1$:

– Gödel's original $\mathcal{IT}_1$ (Theorem 9) and an $\mathcal{IT}_1$ based on classical logic (Theorem 12) required the formalization of some well-known arguments without change.
– Rosser's $\mathcal{IT}_1$ (Theorem 10) involved the generalization of a well-known argument: distilling two abstract conditions, Ord_1 and Ord_2.
– Novel semantic variants of $\mathcal{IT}_1$ (Theorems 11 and 13) were born from abstractly connecting standard models, HBL_1's "iff" version, and proof representability.

They also include two versions of $\mathcal{IT}_2$:

– The standard $\mathcal{IT}_2$ based on the three derivability conditions (Theorem 14) again only required formalizing a well-known argument.
– The alternative, Jeroslow-style $\mathcal{IT}_2$ (Theorems 17 and 18) involved a detailed analysis and correction of an existing abstract result.

12000 LOC

From Abstract to Concrete

Generic Syntax	Connectives	Provability Relation	Numerals

Verified instances

- Robinson’s Arithmetic (Q)
- Hereditarily finite set theory

From Abstract to Concrete

Instantiations of

classical
semantic

with Paulson's
HF set theory.

From Abstract to Concrete

Instantiations of

classical
semantic

with Paulson's
HF set theory.

12000 LOC

From Abstract to Concrete

Instantiations of

classical

semantic

12000 LOC

with Paulson's HF set theory.

Paulson assumes soundness (and redundantly consistency!)

We removed the soundness assumption from the instantiation of 

→ strictly stronger result

→ required us to replace "easy" semantic proofs with tedious proofs in the HF calculus (no help from the abstract side here)

From Abstract to Concrete

Instantiations of

classical

semantic

12000 LOC

with Paulson's HF set theory.

Paulson assumes soundness (and redundantly consistency!)

We removed the soundn… …tion from the instantiation of 

-5000 LOC

→ strictly stronger resul…

→ required us to replace "easy" semantic proofs with tedious proofs in the HF calculus (no help from the abstract side here)

From Abstract to Concrete

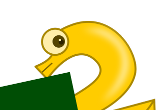
Instantiations of

classical
semantic

12000 LOC

with Paulson's HF set theory.

Paulson assumes soundness (and redundantly consistency!)



We removed the soundn … n from the i…

-5000 LOC

+5000 LOC

→ strictly stronger resul…

→ required us to replace "easy" semantic proofs with tedious proofs in the HF calculus (no help from the abstract side here)

Conclusion

- Abstract formalization of 1 and 2
  - **Answer** "What must/may a logic/theory offer?"
  - **Understand** variants and **distill** trade-offs from the literature
  - **Correct** a mistake in a pen and paper proof
- Concrete instantiation to hereditarily finite set theory
  - **Reproduce** (for 1) and **improve** (for 2) Paulson's formalization
- Still unanswered/future work
  - Do 1 and 2 hold for Intuitionistic FOL, HOL, CIC?
  - Can we do more on the abstract level? (e.g. derivability conditions)

Thank you!
Questions?

Conclusion

- Abstract Isabelle HOL formalization of 1 and 2
  - **Answer** “What must/may a logic/theory offer?”
  - **Understand** variants and **distill** trade-offs from the literature
  - **Correct** a mistake in a pen and paper proof
- Concrete instantiation to hereditarily finite set theory
  - **Reproduce** (for 1) and **improve** (for 2) Paulson’s formalization
- Still unanswered/future work
  - Do 1 and 2 hold for Intuitionistic FOL, HOL, CIC?
  - Can we do more on the abstract level? (e.g. derivability conditions)

A Formally Verified Abstract Account of Gödel's Incompleteness Theorems

Andrei Popescu

Middlesex University London

Dmitriy Traytel

ETH zürich