# Incompleteness v2 Revisited

Harry Sha

Fall 2026

# Section 1

## Review

# Incompleteness v1 (Formal)

Let $\Sigma$ be a set of sentences, and $\mathcal{N}$ be the standard model. Then, the following cannot be satisfied simultaneously.

1. **Completeness**. $\mathcal{N} \models A \implies \Sigma \vdash A$
2. **Soundness**. $\Sigma \vdash A \implies \mathcal{N} \models A$
3. **Recognizable axioms**. $\Sigma \in \mathbf{TR}$
4. **Expressivity**. For each TM $M$, there is a formula $\varphi_M$ such that $\mathcal{N} \models \varphi_M$ if and only if $M$ halts on the emptystring.

We showed that the last property could be removed so long as we are working over the language of arithmetic since were were able to construct this formula $\varphi_M$ by expressing TM configurations (and sequences of them) as numbers, and the transitions of the TM as operations of arithmetic.

# Some formulas we constructed

$\ulcorner w \urcorner$: encoding of $w \in \mathbb{N}$.

- $\mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, h)$. $h$ codes an accepting sequence of TM configurations for $M$ when run on $w$.
- $\mathrm{RejHist}_M(\overline{\ulcorner w \urcorner}, h)$. $h$ codes a rejecting sequence of TM configurations for $M$ when run on $w$.
- $\mathrm{HaltHist}_M(\overline{\ulcorner w \urcorner}, h)$. $h$ codes a halting sequence of TM configurations for $M$ when run on $w$.
- $A_M(\overline{\ulcorner x \urcorner})$. $\exists h \mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, h)$
- $H_M(\overline{\ulcorner x \urcorner})$. $\exists h \mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, h)$

Sometimes we included $t$, the number of steps as an extra parameter, but we can always just pair it with $h$...

# Incompleteness v1 (Formal, stronger)

Let $\Sigma$ be a set of sentences in the language of arithmetic, and $\mathcal{N}$ be the standard model. Then, the following cannot be satisfied simultaneously.

1. **Completeness**. For all $A$, $\Sigma \vdash A$ or $\Sigma \vdash \neg A$
2. **Soundness**. $\mathcal{N} \models \Sigma$
3. **Recognizable axioms**. $\Sigma \in \mathbf{TR}$

Is there an effective axiomatization that exactly captures $\mathrm{Th}(\mathcal{N})$?

$$\mathrm{Th}(\mathcal{N}) = \{A : \mathcal{N} \models A\}.$$

we hope: there is $\Sigma$ s.t. $\Sigma \in TR$, and

$$\mathrm{Thm}(\Sigma) = \mathrm{Th}(\mathcal{N})$$

since $\mathcal{N} \models A$ or $\mathcal{N} \models \neg A$, and not both,

$\Rightarrow \Sigma$ proves exactly one of $A, \neg A$.

$\mathcal{N} \models \Sigma$? Yes b/c $A \in \Sigma \Rightarrow \Sigma \vdash A$.

# Incompleteness v2 (Informal)

No formal system can simultaneously have all four properties:

- **Completeness**. for all $\varphi$ either $\varphi$ or $\neg\varphi$ is provable.
- **Consistency**. for all $\varphi$ at most one of $\varphi$ and $\neg\varphi$ is provable.
- **Verifiable proofs**. there is an algorithm that checks proposed proofs.
- **Expressivity**. the system represents TM computations.

# Incompleteness v2 (Formal)

Let $\Sigma$ be a set of sentences in the language of arithmetic. Then, the following cannot be satisfied simultaneously.

- **Completeness.**. For all $A$, $\Sigma \vdash A$ or $\Sigma \vdash \neg A$
- **Consistency.** For all $A$, at most one of the following $\Sigma \vdash A$, $\Sigma \vdash \neg A$.
- **Recognizable axioms.** $\Sigma \in \mathbf{TR}$
- **Expressivity**. There is a formula $\alpha_x$ such that
  - If $M_x$ accepts $\epsilon$, then $\Sigma \vdash \alpha_x$.
  - If $M_x$ rejects $\epsilon$, then $\Sigma \vdash \neg\alpha_x$.
  - The map $x \mapsto \alpha_x$ is computable.

$\forall u$ s.t.
$u$ halts $\Leftrightarrow \mathcal{N} \models \forall u$.

Note. No mention of the standard model!

# Review of Proof

$R$:

- $\langle R \rangle =$ `get_source()`
- construct $\alpha_{\langle R \rangle}$.
- For each string $p$ in standard string order:
  - if $p$ is a proof of $\alpha_{\langle R \rangle}$, reject;
  - if $p$ is a proof of $\neg\alpha_{\langle R \rangle}$, accept.

Then $\Sigma$ either proves $\alpha_{\langle R \rangle}$ or $\neg\alpha_{\langle R \rangle}$ by completeness, but both cases result in contradiction. (why?)

- If $M_x$ accepts $\epsilon$, then $\Sigma \vdash \alpha_x$.
- If $M_x$ rejects $\epsilon$, then $\Sigma \vdash \neg\alpha_x$.

if $\Sigma \vdash \alpha_R \Rightarrow$ R rejects $\epsilon \Rightarrow \Sigma \vdash \neg\alpha_R$.

other case is similar!

if $\Sigma \vdash \neg\alpha_R \Rightarrow$ R accepts $\epsilon \Rightarrow \Sigma \vdash \alpha_R$

violating consistency!

Just like last lecture, we're going to do away with the expressivity condition. The result will be an even stronger version of the incompleteness theorem.

# Incompleteness v2

Let $\Sigma$ be a set of sentences in the language of arithmetic. The following cannot simulatneously hold. Then, the following cannot be satisfied simultaneously.

- **Completeness.** For all $A$, $\Sigma \vdash A$ or $\Sigma \vdash \neg A$
- **Consistency.** For all $A$, at most one of the following $\Sigma \vdash A$, $\Sigma \vdash \neg A$.
- **Recognizable axioms.** $\Sigma \in \mathbf{TR}$
- **Captures basic arithmetic.** $\mathrm{Cn}(A_E) \subseteq \mathrm{Cn}(\Sigma)$

I haven't yet told you what $A_E$ is, so let's start with that.

**Exercise.** Show that the "captures basic arithmetic condition" is implied by the Completeness and Soundness conditions from Incompleteness v1.

# Section 2

$A_E$

# The set of axioms $A_E$

$$\forall x \quad \mathbf{S}x \neq \mathbf{0} \tag{S1}$$

$$\forall x \forall y \quad (\mathbf{S}x = \mathbf{S}y \rightarrow x = y) \tag{S2}$$

$$\forall x \forall y \quad (x < \mathbf{S}y \leftrightarrow x \leq y) \tag{L1}$$

$$\forall x \quad x \not< \mathbf{0} \tag{L2}$$

$$\forall x \forall y \quad (x < y \vee x = y \vee y < x) \tag{L3}$$

$$\forall x \quad x + \mathbf{0} = x \tag{A1}$$

$$\forall x \forall y \quad x + \mathbf{S}y = \mathbf{S}(x + y) \tag{A2}$$

$$\forall x \quad x \cdot \mathbf{0} = \mathbf{0} \tag{M1}$$

$$\forall x \forall y \quad x \cdot \mathbf{S}y = x \cdot y + x \tag{M2}$$

$$\forall x \quad x\mathbf{E}\mathbf{0} = \mathbf{S}\mathbf{0} \tag{E1}$$

$$\forall x \forall y \quad x\,\mathbf{E}\,\mathbf{S}y = x\,\mathbf{E}\,y \cdot x \tag{E2}$$

# Obervations

- $A_E$ is finite.
- $\mathcal{N} \models A_E$

Question. What is the relationship between $\mathrm{Th}(\mathcal{N})$ and $\mathrm{Cn}(A_E)$.

$\supseteq$

$\mathrm{Cn}(A_E) \subseteq \mathrm{Th}(\mathcal{N})$

$A_E \vdash \varphi \Rightarrow \varphi$ is true in all models of $A_E$, we have $\mathcal{N} \models A_E$. Hence $\mathcal{N} \models \varphi$.

# Review of the Goal

It turns out, $A_E$ is already strong enough so that we can construct $\alpha_x$ with the properties that we wanted:

- If $M_x$ accepts $\epsilon$, then $A_E \vdash \alpha_x$.
- If $M_x$ rejects $\epsilon$, then $A_E \vdash \neg\alpha_x$.

# So what kinds of things can $A_E$ prove?

Recall $\overline{n}$ is short for $S...S0$ where there are $n$ $S$s.

In the pre-lecture questions, I asked you to prove $\overline{1} + \overline{1} = \overline{2}$.

For any natural number $k$, $A_E \vdash \forall x(x < \overline{k+1} \leftrightarrow x = \overline{0} \vee \ldots \vee x = \overline{k})$

By induction

Base case: $k=0$. WTS $x < \overline{1} \leftrightarrow x = \overline{0}$

$x < \overline{1} \leftrightarrow x \leq 0 \quad (L_1)$

$\leftrightarrow x = 0 \vee x < 0 \quad (\leq)$

$\leftrightarrow x = 0 \quad (L_2)$

Inductive: Assume claim for $k$, Let's show for $k+1$.

$x < \overline{k+2} \leftrightarrow x \leq \overline{k+1} \quad (L_1)$

$\leftrightarrow x = \overline{k+1} \vee x < \overline{k+1} \quad (\leq)$ $\leftrightarrow$ $x = \overline{k+1} \vee x = \overline{0} \vee \ldots \vee x = \overline{k} \quad (IH)$

For all terms $t$ without any variables, $A_E \vdash t = \overline{k}$ for some $k \in \mathbb{N}$.

Induction on terms.

Base case: $t = 0$. $A_E \vdash 0 = 0$ (equality) ✓

Inductive step, assume $A_E \vdash t = \overline{k}$ and $A_E \vdash s = \overline{j}$

WTS the claim for $S(t)$, $t + s$, $t \cdot s$, $E(t, s)$. $t + s = \overline{k} + \overline{j}$ $(=)$

$S(t)$: By IH $A_E \vdash t = \overline{k}$

$S(t) = S(\overline{k})$ $(\text{by } =)$

$= \overline{k+1}$

$= S(\overline{k} + \overline{j-1})$ $(A_2)$

$j$ ⋮

$= S^j(\overline{k} + \overline{0})$

$= S^j(\overline{k})$ $(A_1)$

$= \overline{k+j}$

For any quantifier-free formula $A$ such that $\mathcal{N} \models A$, $A_E \vdash A$.

By induction on formulas.

Base case $t = s$ or $t < s$.

$\mathcal{N} \models t = s$, $A_E \vdash t = \bar{k}$ $A_E \vdash s = \bar{j}$

$\Rightarrow k = j$

so. $A_E \vdash \bar{k} = \bar{j}$

$A_E \vdash t = s$

$\mathcal{N} \models t < s$. let say $t^{\mathcal{N}} = k$, $s^{\mathcal{N}} = j$

$t^{\mathcal{N}} < s^{\mathcal{N}}$

$k < j$.

$A_E \vdash \bar{k} < \bar{j}$ by $L_1$ (many times).

$A_E \vdash t < s$. $(=)$

Define an existential formula as a formula of the form $\exists x_1 ... \exists x_k B$ where $B$ is quantifier-free

For any existential formula $A$ such that $\mathcal{N} \models A$, $A_E \vdash A$.

$A = \exists x B$ then, $\mathcal{N} \models \exists x B(x)$

$\Leftrightarrow \mathcal{N}, s[x \to a] \models B(x)$ for some $a \in \mathcal{N}$.

$\Leftrightarrow \mathcal{N} \models B(\bar{a})$ (substitution lemma)

$\Leftrightarrow A_E \vdash B(\bar{a})$ ($B(\bar{a})$ is quantifier free).

$\Leftarrow A_E \vdash \exists x B(x)$ (found witness $\bar{a}$).

Define a bounded formula as a formula in which all quantifiers appear in the following way.

$$\forall x(x < t \to \ldots), \exists x(x < t \wedge \ldots)$$

For any bounded formula $A$ such that $\mathcal{N} \models A$, $A_E \vdash A$.

in fact $A_E \vdash \forall x < \bar{n}.\, B(x) \iff B(\bar{0}) \wedge B(\bar{1}) \wedge \ldots \wedge B(\overline{n-1})$

suppose $A_E \vdash \forall x\,(x < \bar{n} \to B(x)) \to A$

$\vdash \forall x\,(x = \bar{0} \vee x = \bar{1} \vee \ldots \vee x = \overline{n-1} \to B(x))$.

since $\mathcal{N} \models A$. then for $m < n$, $\mathcal{N} \models B(\bar{m}) \Rightarrow A_E \vdash B(\bar{m})$.

By instantiation $A_E \vdash B(\bar{m})$, can do this for all $m < n$.

# The formulas from last time

- $\mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, h)$. $h$ codes an accepting sequence of TM configurations for $M$ when run on $w$.
- $\mathrm{RejHist}_M(\overline{\ulcorner w \urcorner}, h)$. $h$ codes a rejecting sequence of TM configurations for $M$ when run on $w$.

Note these are both bounded!

# Attempt

$M$ accepts $\Rightarrow A_E \vdash \alpha_x$

$M$ rejects $\Rightarrow A_E \vdash \neg\alpha_x$.

Take $\alpha_x = \exists h.\mathrm{AccHist}_{M_x}(\overline{\ulcorner\epsilon\urcorner}, h)$?

if $M$ accepts $\Rightarrow \mathcal{N} \models \alpha_x \Rightarrow A_E \vdash \alpha_x$ ✓

last lecture

↑ $\alpha_x$ true in $\mathcal{N}$, and an existential of a bounded formula.

if $M$ rejects we want $A_E \vdash \forall h\, \neg\mathrm{AccHist}_{M_x}(\overline{\ulcorner\varepsilon\urcorner}, h)$.

So $\mathcal{N} \models \neg\alpha$, we can prove $\neg\mathrm{AccHist}_{M_x}(\overline{\ulcorner\varepsilon\urcorner}, \overline{0})$

$\neg\mathrm{AccHist}_{M_x}(\overline{\ulcorner\varepsilon\urcorner}, \overline{1})$ ...

# Aside: Non Standard Models

construct $\mathcal{U}$ st $\mathrm{Th}(\mathcal{U}) = \mathrm{Th}(\mathcal{N})$

To illustrate the previous problem. Consider the following construction of a model of $\mathrm{Th}(\mathcal{N})$. Add a constant $c$, and consider the set.

Take $T = \mathrm{Th}(\mathcal{N}) \cup \{\overline{n} < c : n \in \mathbb{N}\}$

- Let $T_0$ be any finite subset of $T$.
- $\mathcal{N} \models T_0$

$\Longrightarrow$ Compactness $T$ is satisfiable, lets say by $\mathcal{U}$. then $c$ cannot be assigned to any natural number.

# The initial proof

Gödel's initial proof of the incompleteness theorem sidestepped this problem using a stronger assumption. Namely, he assumed something called $\omega$-consistency which says that if the system can prove $P(\overline{0}), P(\overline{1}), ...$, then you can't prove $\exists x \neg P(x)$.

Combined with the completeness assumption, it says if you can prove $P(\overline{0}), P(\overline{1})$,..., then you can also prove $\forall x P(x)$.

Later, Rosser showed that this assumption is not needed.

# Rosser's Trick

For every TM $M$, and input $w$, define

$$\Phi_M(\overline{\ulcorner w \urcorner}) = \exists a.(\mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, a) \wedge \forall r < a.\neg \mathrm{RejRun}_M(\overline{\ulcorner w \urcorner}, r))$$

## Lemma (Smaller opposing witness)

*For every TM $M$, and input string $w$.*

- *$M$ accepts $w$ implies $A_E \vdash \Phi_M(\overline{\ulcorner w \urcorner})$.*
- *$M$ rejects $w$ implies $A_E \vdash \neg\Phi_M(\overline{\ulcorner w \urcorner})$.*
- *$\Phi_M$ is computable given a descprition of $M$.*

In otherwords, $\Phi_{M_x}(\overline{\ulcorner \epsilon \urcorner})$ is the formula, we want.

# Proof

$$\Phi_M(\overline{\ulcorner w \urcorner}) = \exists a.(\mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, a) \wedge \forall r < a.\neg \mathrm{RejRun}_M(\overline{\ulcorner w \urcorner}, r))$$

$M$ accepts $w$. $\Rightarrow$ $\Phi_M$ is true in $\mathcal{N}$, $\Phi$ is $\exists a$ Banded, so

$$A_E \vdash \Phi_M.$$

$M$ reject $w$. $\Rightarrow$ then let $h$ be the code for a rejecting history.

WTS: $A_E \vdash \neg \Phi_M$

i.e. $A_E \vdash \forall a \left(\neg \mathrm{AccHist}_M(\overline{\ulcorner w \urcorner}, a) \vee \exists r < a \ \mathrm{RejRun}(\overline{\ulcorner w \urcorner}, r)\right)$

$A_E \vdash \forall a \left( \neg \mathrm{AccHist}_M(\overline{\lceil w \rceil}, a) \vee \exists r < a \ \mathrm{RejRun}(\overline{\lceil w \rceil}, r) \right)$

Let $a$ be arbitrary, we have by $(L_3)$,

$$a \geq h+1 \quad \text{or} \quad a \leq h$$

if $a \geq h+1$, then there is an $r$, namely $h$,
s.t. $r < a$ and $\mathrm{RejRun}(\overline{\lceil w \rceil}, r)$.
hence $A_E \vdash \exists r < a \ \mathrm{RejRun}(\lceil w \rceil, r)$.

if $a \leq h$, $A_E \vdash \neg \underbrace{\mathrm{AccHist}_M(\lceil w \rceil, \bar{a})}$

We free & bounded.

# Section 3

## You don't need E

We can even remove $E$ and the related axioms from $A_E$, and all the proofs from before still work. In particular, we can simulate $E$ with just multiplication and addition.

# Sketch

Let $A_M \subseteq A_E$ be the subset without the exponentiation axioms. We'll show that there is a formula $\mathrm{Exp}(x, y, z)$ not containing $E$ such that

- $A_M \vdash \mathrm{Exp}(\overline{m}, \overline{n}, \overline{m^n})$.
- $A_M \vdash \forall z \mathrm{Exp}(\overline{m}, \overline{n}, z) \to z = \overline{m^n}$.

An initial idea is just to set $\mathrm{Exp}(x, y, z) = z = x...x$, but this doesn't work...

# Sketch

See Enderton for the details!

# Section 4

# Second Incompleteness

# The Second Incompleteness Theorem (Informal)

Suppose $\Sigma$ satisfies the following

- **Consistent**. For all sentences $A$, it's not the case that $\Sigma \vdash A$ and $\Sigma \vdash \neg A$.
- $\Sigma \in \mathbf{TR}$.
- **Captures basic arithmetic.** $\mathrm{Cn}(A_E) \subseteq \mathrm{Cn}(\Sigma)$,
- $\Sigma$ also satisfies some slightly stronger conditions.

Then $\Sigma \nvdash \mathrm{Con}(\Sigma)$

# $\mathsf{Con}(\Sigma)$

$\Sigma$ being consistent in the statement of the theorem is a meta-level definition. However, if we want to say something about $\Sigma$ not proving its own consistency, we must formulate being consistent as a FOL sentence.

# The slightly stronger conditions

1. If $C(a)$ is bounded, and $\sigma = \exists a(C(a))$, then $\Sigma \vdash \sigma \to \mathrm{Prov}_\Sigma(\overline{\ulcorner \sigma \urcorner})$
2. If a TM $M$ enters an accepting state after its proof checker verifies a proof of a formula $A$, then $\Sigma \vdash \mathrm{Accept}_M(\overline{\ulcorner w \urcorner}) \to \mathrm{Prov}_\Sigma(\overline{\ulcorner A \urcorner})$
3. $\Sigma \vdash \mathrm{Prov}_\Sigma(\overline{\ulcorner A \urcorner}) \wedge \mathrm{Prov}_\Sigma(\overline{\ulcorner \neg A \urcorner}) \to \mathrm{Prov}_\Sigma(\overline{\ulcorner A \wedge \neg A \urcorner})$

# Proof

Use the same program as before.

$R$:

- $\langle R \rangle = \texttt{get_source()}$
- construct $\alpha_{\langle R \rangle}$.
- For each string $p$ in standard string order:
  - if $p$ is a proof of $\alpha_{\langle R \rangle}$, reject;
  - if $p$ is a proof of $\neg\alpha_{\langle R \rangle}$, accept.

Recall, we ended up choosing
$\alpha_{\langle R \rangle} = \Phi_{\langle R \rangle}(\overline{\ulcorner w \urcorner}) = \exists a.(\mathrm{AccHist}_R(\overline{\ulcorner w \urcorner}, a) \wedge \forall r < a.\neg\mathrm{RejRun}_R(\overline{\ulcorner w \urcorner}, r))$

For simplicity, let's abbreviate that entire formula as $A$

# Lemma: $\Sigma \vdash A \rightarrow \neg \mathsf{Con}(A)$

$A = \exists a.(\mathrm{AccHist}_R(\overline{\ulcorner w \urcorner}, a) \wedge \forall r < a.\neg \mathrm{RejRun}_R(\overline{\ulcorner w \urcorner}, r))$

# Conclusion