Gödel’s Second Incompleteness Theorem

**Set (consistency):** A proof system \( \mathcal{P} \) is consistent if \( \mathcal{P} \) doesn’t prove (derive) both \( A \) and \( \neg A \) for some \( A \).

A proof system \( \mathcal{P} \) for arithmetic is consistent if \( \mathcal{P} \) doesn’t prove "1=2" (\( \mathcal{P} \) will prove "1≠2" from the Peano arithmetic axioms).

\[
\text{Cons} \quad \equiv \quad " \ (1=2) \text{ is not provable in } \mathcal{P} "
\]

\( \mathcal{P} \)

\[\uparrow\]

an arithmetic formula expressing that \( \mathcal{P} \) is consistent.

(We’ve seen before that provability in \( \mathcal{P} \) can be expressed by an arithmetic formula.)

\[\text{Consistent} \quad \frac{\text{syntactic}}{\text{semantic}} \quad \text{Sound} \quad \frac{\text{syntactic}}{\text{semantic}} \]

\[\text{Sound} \quad \frac{\text{syntactic}}{\text{semantic}} \]

\[\text{Not Sound} \quad \frac{\text{syntactic}}{\text{semantic}} \]

\[\text{Not Consistent} \quad \frac{\text{syntactic}}{\text{semantic}} \]
**Syntax**

\[ \text{Cons}_S \equiv \neg \exists S \vdash "1=2" \]

... doesn't prove (derive)

**Semantics**

\[ \forall \psi [(S \vdash \psi) \Rightarrow \text{True}(\psi)] \]

there is no ar. formula expressing the truth [Tarski's Thm (see later)]

**Theorem** (Gödel's 2nd Incompleteness)

Fix any proof system $S$ powerful enough to reason about $+,\times$ & also satisfying some provability conditions.

If $S$ is consistent, then

\[ S \not\vdash \text{Cons}_S \]

(i.e., $S$ cannot prove its own consistency)

**Claim:** If $S$ is consistent, then

\[ G \equiv " I'm not provable in $S" \]

(considered last time)
is not provable in $\mathcal{P}$.

Proof:
Let $g = \varphi(g)$ (the Gödel number of $g$)

Then

$$g \equiv \neg \exists x \text{ Proof}(x, [g])$$

Suppose $g$ is provable in $\mathcal{P}$, i.e., $\mathcal{P} \vdash g$.

Then $\exists m \in \mathbb{N}$ s.t.

$$\text{Proof}(\left[ m \right], [g])$$

is true (over $\mathbb{N}$).

By provability assumptions on $\mathcal{P}$, we get

$$\mathcal{P} \vdash \text{Proof}(\left[ m \right], [g])$$

(intuitively, $\exists$ an algorithm to check if $\text{Proof}(a,b)$ is True, for any given input numbers $a,b \in \mathbb{N}$.
$\mathcal{P}$ can simulate this algorithm on $\left[ m \right]$ and $[g]$. )
on Lmod and Lg  

Then we get that

\[ \exists x \text{ Proof}(x, \mathcal{L}_g) \]

\[ \Downarrow \]

\[ \forall \mathcal{L}_g \]

So, \[ \forall \mathcal{L}_g \]

But (earlier we assumed also) \[ \exists \mathcal{L}_g \]

\[ \Downarrow \]

\[ \forall \mathcal{L}_g (\mathcal{L}_g \land \neg \mathcal{L}_g) \]

\[ \Rightarrow \]

\[ \forall \mathcal{L}_g \quad "1 = 2" \]

Thus, \( \forall \mathcal{L}_g \) is inconsistent.

We conclude: \( \text{Cons} \mathcal{L}_g \rightarrow \mathcal{L}_g \)

is a true implication.
Proof of Gödel's Theorem:

The proof of claim above can be formalized within $\mathcal{P}$ itself!

So,

$$\mathcal{P} \vdash (\text{Cons}_\mathcal{P} \rightarrow \mathcal{G})$$

Suppose $\mathcal{P} \vdash \text{Cons}_\mathcal{P}$

Then

$$\mathcal{P} \vdash \mathcal{G}$$

Moreover (by provability assumptions on $\mathcal{P}$)

$$\mathcal{P} \vdash "\mathcal{P} \vdash \mathcal{G}"$$

$$\mathcal{P} \vdash \mathcal{G}$$
But then $\Phi \vdash \phi \land \neg \phi$, so $\Phi$ is inconsistent. Hence, consistent $\Phi$ cannot prove $\neg \neg \Phi$. \[\square\]

---

**Computability**

1. $\exists$ undecidable problems (non-constructive)

2. Example $\exists$ hard language (not semi-dec.)

3. Natural problems like Halting that are hard.

---

**Logic**

1. A sound $\Phi$

2. Example $\exists$ is true but not provable in any sound $\Phi$.

3. Cons $\Phi$ is true but not provable in consistent $\Phi$. 

4. O ...
4. Recursion Theorem

Thm ("Recursion theorem for ar. formulas")

Let $A(x)$ be any ar. formula. Then $\exists$ ar. sentence $B$ s.t.

$$B = A([n_0])$$

where $n_0 = \gamma(B)$ is the Gödel number of $B$.

Proof: Using sub (Sub)

$$C(x) := A(Sub(x,x))$$

Let $a = \gamma(C(x))$.

Define $B = C([Ca])$. 
Define \( B = C(L_{aJ}) \)

\[ \equiv A \left( \text{Sub}([a], [a]) \right) \]

\[ \forall (B) = \text{sub}([a], [a]) \forall (B) \]


Tarski’s Thm: There is no ar. formula \( \text{Truth}(x) \)

s.t.

\( \forall n \)

\( \text{Truth}([n]) \) holds if

sentence \( [n] \) is true over \( \mathbb{N} \).

Proof: Suppose \( \text{Truth}(x) \) has ar. form.

\[ A(x) = \neg \text{Truth}(x) \]

By Rec. Thm for Ar. Form’s.
\[ f' \]
\[ B = ? \text{Truth}(\phi(B)) \]
\[ = "I am false" \]