Lecture 17

Wednesday, June 19, 2019 11:59 AM

Gödel's Second Incompleteness Theorem

<u>Def</u> (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 "$1 \neq 2$" from the Peano arithm. axioms).

$$Cons_{\mathcal{P}} \equiv \text{" } (1=2) \text{ is not provable in } \mathcal{P} \text{ "}$$

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

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

Consistent	Sound
• syntactic	• semantic

- syntactic
- $\exists$ arithm. formula

$$\mathrm{Cons}_{\mathfrak{P}} \equiv [\mathfrak{P} \nvdash \text{"}1=2\text{"}]$$

doesn't prove (derive)

- semantic
- $\forall \psi \, [(\mathfrak{P} \vdash \psi) \Rightarrow \mathrm{True}(\psi)]$

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

Thm (Gödel's 2nd Incompleteness)

Fix any proof system $\mathfrak{P}$ powerful enough to reason about $+$, $*$ & also satisfying some provability conditions.

If $\mathfrak{P}$ is consistent, then

$$\mathfrak{P} \nvdash \mathrm{Cons}_{\mathfrak{P}}$$

(i.e., $\mathfrak{P}$ cannot prove its own consistency).

Claim: If $\mathfrak{P}$ is consistent, then

$G \equiv$ "I'm not provable in $\mathfrak{P}$"

(considered last time)

(considered last time)'
is not provable in $\mathfrak{T}$.

<u>Proof</u>:

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

Then

$$G \equiv \neg \exists x \, \mathrm{Proof}(x, [g]).$$

Suppose G is provable in $\mathfrak{T}$, i.e., $\mathfrak{T} \vdash G$.

Then $\exists \, m_0 \in \mathbb{N}$ s.t.

$$\mathrm{Proof}([m_0], [g])$$

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

By provability assumptions on $\mathfrak{T}$, we get

$$\mathfrak{T} \vdash \mathrm{Proof}([m_0], [g])$$

(intuitively, $\exists$ an algorithm to check if $\mathrm{Proof}(a, b)$ is True, for any given input numbers $a, b \in \mathbb{N}$. $\mathfrak{T}$ can simulate this algorithm on $[m_0]$ and $[g]$.)

on $\ulcorner m_0 \urcorner$ and $\ulcorner g \urcorner$.)

Then we get that

$$\mathfrak{T} \vdash \underbrace{\exists x\, \mathrm{Proof}(x, \ulcorner g \urcorner)}_{\substack{\equiv \\ \neg G}}$$

So, $\mathfrak{T} \vdash \neg G$

But (earlier we assumed also) $\mathfrak{T} \vdash G$

$\Downarrow$

$$\mathfrak{T} \vdash (G \wedge \neg G)$$

$\Rightarrow$

$$\mathfrak{T} \vdash \text{"} 1 = 2 \text{"}$$

Thus, $\mathfrak{T}$ is inconsistent.

We conclude: $\mathrm{Cons}_{\mathfrak{T}} \to G$

is a true implication. $\square$

Proof of Gödel's Theorem:

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

So,

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

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

Then

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

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

$$\mathcal{P} \vdash \text{“}\underbrace{\mathcal{P} \vdash G}_{\equiv\ \neg G}\text{”}$$

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

But then $\mathcal{J} \vdash G \wedge \neg G$,
so $\mathcal{J}$ is inconsistent.
Hence, consistent $\mathcal{J}$ cannot prove $Cons_{\mathcal{J}}$. $\square$

Computability	Logic
1. $\exists$ undecidable problems (non-constructive)	1. $\forall$ sound $\mathcal{J}$ $\exists$ true but not provable sentence (non-constructive)
2. Example D hard language (not semi-dec.)	2. Example G is true but not provable in any sound $\mathcal{J}$.
3. Natural problems like Halting that are hard.	3. $Cons_{\mathcal{J}}$ is true but not provable in consistent $\mathcal{J}$.

consistent ?

4. Recursion Thm 4.

Thm ("Recursion thm for ar. Formulas")

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

$$B \equiv A([n_0])$$

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

Proof: Using sub (Sub)

$$C(x) := A(\mathrm{Sub}(x,x))$$

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

Define $B = C([a])$.

Define $B = C(\ulcorner a \urcorner)$

$$\equiv A(\underbrace{Sub(\ulcorner a \urcorner, \ulcorner a \urcorner)}_{= \varphi(B)})$$

$$\varphi(B) = sub(\ulcorner a \urcorner, \ulcorner a \urcorner)$$

$\square$

Tarski's Thm: There is no ar. formula $Truth(x)$ s.t. $\forall n$

$Truth(\ulcorner n \urcorner)$ holds $\Leftrightarrow$ sentence $\ulcorner n \urcorner$ is true over $\mathbb{N}$.

Proof: Suppose $Truth(x)$ has ar. form.

$$\boxed{A(x) = \neg \underline{Truth(x)}}$$

By Rec Thm for Ar. Form's,

$\neg$

$\exists'$

$$B \equiv \neg \mathrm{Truth}(\varphi(B))$$

$$\equiv \text{"I am false"}$$