

inp.1 The Second Incompleteness Theorem

inc:inp:2in:
sec

How can we express the assertion that **PA** doesn't prove its own consistency? Saying **PA** is inconsistent amounts to saying that **PA** proves $0 = 1$. So we can take $\text{Con}_{\mathbf{PA}}$ to be the formula $\neg\text{Prov}_{\mathbf{PA}}(\ulcorner 0 = 1 \urcorner)$, and then the following theorem does the job:

inc:inp:2in:
thm:second-incompleteness

Theorem inp.1. *Assuming **PA** is consistent, then **PA** does not prove $\text{Con}_{\mathbf{PA}}$.*

It is important to note that the theorem depends on the particular representation of $\text{Con}_{\mathbf{PA}}$ (i.e., the particular representation of $\text{Prov}_{\mathbf{PA}}(y)$). All we will use is that the representation of $\text{Prov}_{\mathbf{PA}}(y)$ has the three properties above, so the theorem generalizes to any theory with a provability predicate having these properties.

It is informative to read Gödel's sketch of an argument, since the theorem follows like a good punch line. It goes like this. Let $\gamma_{\mathbf{PA}}$ be the Gödel sentence that we constructed in the proof of ???. We have shown "If **PA** is consistent, then **PA** does not prove $\gamma_{\mathbf{PA}}$." If we formalize this *in* **PA**, we have a proof of

$$\text{Con}_{\mathbf{PA}} \rightarrow \neg\text{Prov}_{\mathbf{PA}}(\ulcorner \gamma_{\mathbf{PA}} \urcorner).$$

Now suppose **PA** proves $\text{Con}_{\mathbf{PA}}$. Then it proves $\neg\text{Prov}_{\mathbf{PA}}(\ulcorner \gamma_{\mathbf{PA}} \urcorner)$. But since $\gamma_{\mathbf{PA}}$ is a Gödel sentence, this is equivalent to $\gamma_{\mathbf{PA}}$. So **PA** proves $\gamma_{\mathbf{PA}}$.

But: we know that if **PA** is consistent, it doesn't prove $\gamma_{\mathbf{PA}}$! So if **PA** is consistent, it can't prove $\text{Con}_{\mathbf{PA}}$.

To make the argument more precise, we will let $\gamma_{\mathbf{PA}}$ be the Gödel sentence for **PA** and use the provability conditions (1)–(3) above to show that **PA** proves $\text{Con}_{\mathbf{PA}} \rightarrow \gamma_{\mathbf{PA}}$. This will show that **PA** doesn't prove $\text{Con}_{\mathbf{PA}}$. Here is a sketch

of the proof, in **PA**. (For simplicity, we drop the **PA** subscripts.)

$$\begin{aligned}
\gamma &\leftrightarrow \neg\text{Prov}(\ulcorner\gamma\urcorner) & (1) & \text{inc:inp:2in:} \\
&\quad \gamma \text{ is a Gödel sentence} & & \text{G2-1} \\
\gamma &\rightarrow \neg\text{Prov}(\ulcorner\gamma\urcorner) & (2) & \text{inc:inp:2in:} \\
&\quad \text{from eq. (1)} & & \text{G2-2} \\
\gamma &\rightarrow (\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \perp) & (3) & \text{inc:inp:2in:} \\
&\quad \text{from eq. (2) by logic} & & \text{G2-3} \\
\text{Prov}(\ulcorner\gamma \rightarrow (\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \perp)\urcorner) & & (4) & \text{inc:inp:2in:} \\
&\quad \text{by from eq. (3) by condition P1} & & \text{G2-4} \\
\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \text{Prov}(\ulcorner(\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \perp)\urcorner) & & (5) & \text{inc:inp:2in:} \\
&\quad \text{from eq. (4) by condition P2} & & \text{G2-5} \\
\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow (\text{Prov}(\ulcorner\text{Prov}(\ulcorner\gamma\urcorner)\urcorner) \rightarrow \text{Prov}(\ulcorner\perp\urcorner)) & & (6) & \text{inc:inp:2in:} \\
&\quad \text{from eq. (5) by condition P2 and logic} & & \text{G2-6} \\
\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \text{Prov}(\ulcorner\text{Prov}(\ulcorner\gamma\urcorner)\urcorner) & & (7) & \text{inc:inp:2in:} \\
&\quad \text{by P3} & & \text{G2-7} \\
\text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \text{Prov}(\ulcorner\perp\urcorner) & & (8) & \text{inc:inp:2in:} \\
&\quad \text{from eq. (6) and eq. (7) by logic} & & \text{G2-8} \\
\text{Con} \rightarrow \neg\text{Prov}(\ulcorner\gamma\urcorner) & & (9) & \text{inc:inp:2in:} \\
&\quad \text{contraposition of eq. (8) and } \text{Con} \equiv \neg\text{Prov}(\ulcorner\perp\urcorner) & & \text{G2-9} \\
\text{Con} \rightarrow \gamma & & & \\
&\quad \text{from eq. (1) and eq. (9) by logic} & &
\end{aligned}$$

The use of logic in the above just elementary facts from propositional logic, e.g., eq. (3) uses $\vdash \neg\varphi \leftrightarrow (\varphi \rightarrow \perp)$ and eq. (8) uses $\varphi \rightarrow (\psi \rightarrow \chi), \varphi \rightarrow \psi \vdash \varphi \rightarrow \chi$. The use of condition P2 in eq. (5) and eq. (6) relies on instances of P2, $\text{Prov}(\ulcorner\varphi \rightarrow \psi\urcorner) \rightarrow (\text{Prov}(\ulcorner\varphi\urcorner) \rightarrow \text{Prov}(\ulcorner\psi\urcorner))$. In the first one, $\varphi \equiv \gamma$ and $\psi \equiv \text{Prov}(\ulcorner\gamma\urcorner) \rightarrow \perp$; in the second, $\varphi \equiv \text{Prov}(\ulcorner G\urcorner)$ and $\psi \equiv \perp$.

The more abstract version of the incompleteness theorem is as follows:

Theorem inp.2. *Let \mathbf{T} be any axiomatized theory extending \mathbf{Q} and let $\text{Prov}_{\mathbf{T}}(y)$ be any formula satisfying provability conditions P1–P3 for \mathbf{T} . Then if \mathbf{T} is consistent, then \mathbf{T} does not prove $\text{Con}_{\mathbf{T}}$.* inc:inp:2in: thm:second-incompleteness-gen

Problem inp.1. Show that **PA** proves $\gamma_{\text{PA}} \rightarrow \text{Con}_{\text{PA}}$.

digression

The moral of the story is that no “reasonable” consistent theory for mathematics can prove its own consistency. Suppose \mathbf{T} is a theory of mathematics that includes \mathbf{Q} and Hilbert’s “finitary” reasoning (whatever that may be). Then, the whole of \mathbf{T} cannot prove the consistency of \mathbf{T} , and so, a fortiori, the finitary fragment can’t prove the consistency of \mathbf{T} either. In that sense, there cannot be a finitary consistency proof for “all of mathematics.”

There is some leeway in interpreting the term “finitary,” and Gödel, in the 1931 paper, grants the possibility that something we may consider “finitary” may lie outside the kinds of mathematics Hilbert wanted to formalize. But Gödel was being charitable; today, it is hard to see how we might find something that can reasonably be called finitary but is not formalizable in, say, **ZFC**.

Photo Credits

Bibliography