Rosser’s Theorem

Can we modify Gödel’s proof to get a stronger result, replacing “ω-consistent” with simply “consistent”? The answer is “yes,” using a trick discovered by Rosser. Rosser’s trick is to use a “modified” provability predicate $RProv_T(y)$ instead of $Prov_T(y)$.

Theorem inp.1. Let $T$ be any consistent, axiomatizable theory extending $Q$. Then $T$ is not complete.

Proof. Recall that $Prov_T(y)$ is defined as $\exists x \ Prf_T(x,y)$, where $Prf_T(x,y)$ represents the decidable relation which holds if $x$ is the Gödel number of a derivation of the sentence with Gödel number $y$. The relation that holds between $x$ and $y$ if $x$ is the Gödel number of a refutation of the sentence with Gödel number $y$ is also decidable. Let $not(x)$ be the primitive recursive function which does the following: if $x$ is the code of a formula $\varphi$, $not(x)$ is a code of $\neg \varphi$. Then $Ref_T(x,y)$ holds iff $Prf_T(x,not(y))$. Let $Ref_T(x,y)$ represent it. Then, if $T \vdash \neg \varphi$ and $\delta$ is a corresponding derivation, $Q \vdash Ref_T(\delta, \varphi)$. We define $RProv_T(y)$ as

$$\exists x (Prf_T(x,y) \land \forall z (z < x \rightarrow \neg Ref_T(z,y))).$$

Roughly, $RProv_T(y)$ says “there is a proof of $y$ in $T$, and there is no shorter refutation of $y$.” (You might find it convenient to read $RProv_T(y)$ as “$y$ is shmovable.”) Assuming $T$ is consistent, $RProv_T(y)$ is true of the same numbers as $Prov_T(y)$; but from the point of view of provability in $T$ (and we now know that there is a difference between truth and provability!) the two have different properties. (If $T$ is inconsistent, then the two do not hold of the same numbers!)

By the fixed-point lemma, there is a formula $\rho_T$ such that

$$Q \vdash \rho_T \leftrightarrow \neg RProv_T(\varphi_T). \tag{1}$$

In contrast to the proof of ??, here we claim that if $T$ is consistent, $T$ doesn’t prove $\rho_T$, and $T$ also doesn’t prove $\neg \rho_T$. (In other words, we don’t need the assumption of $\omega$-consistency.)

First, let’s show that $T \not\vdash \rho_T$. Suppose it did, so there is a derivation of $\rho_T$ from $T$; let $n$ be its Gödel number. Then $Q \vdash Prf_T(\pi, \rho_T)$, since $Prf_T$ represents $Prf_T$ in $Q$. Also, for each $k < n$, $k$ is not the Gödel number of $\neg \rho_T$, since $T$ is consistent. So for each $k < n$, $Q \vdash \neg Ref_T(\pi, \rho_T)$. By ??(2), $Q \vdash \forall z (z < \pi \rightarrow \neg Ref_T(z, \rho_T))$. Thus,

$$Q \vdash \exists x (Prf_T(x, \rho_T) \land \forall z (z < x \rightarrow \neg Ref_T(z, \rho_T))),$$

but that’s just $RProv_T(\varphi_T)$. By eq. (1), $Q \vdash \neg \rho_T$. Since $T$ extends $Q$, also $T \vdash \neg \rho_T$. We’ve assumed that $T \vdash \rho_T$, so $T$ would be inconsistent, contrary to the assumption of the theorem.
Now, let’s show that $T \not \vdash \neg \rho_T$. Again, suppose it did, and suppose $n$ is the Gödel number of a derivation of $\neg \rho_T$. Then $\text{Ref}_T(n,^\# \rho_T^\#)$ holds, and since $\text{Ref}_T$ represents $\text{Ref}_T$ in $Q$, $Q \vdash \text{Ref}_T(\pi,^\gamma \rho_T^\gamma)$. We’ll again show that $T$ would then be inconsistent because it would also prove $\rho_T$. Since $Q \vdash \rho_T \iff \neg \text{RProv}_T(\gamma \rho_T^{-})$, and since $T$ extends $Q$, it suffices to show that $Q \vdash \neg \text{RProv}_T(\gamma \rho_T^{-})$. The sentence $\neg \text{RProv}_T(\gamma \rho_T^{-})$, i.e.,

$$
\neg \exists x (\text{Prf}_T(x, \gamma \rho_T^{-}) \land \forall z (z < x \rightarrow \neg \text{Ref}_T(z, \gamma \rho_T^{-}))
$$

is logically equivalent to

$$
\forall x (\text{Prf}_T(x, \gamma \rho_T^{-}) \rightarrow \exists z (z < x \land \text{Ref}_T(z, \gamma \rho_T^{-})))
$$

We argue informally using logic, making use of facts about what $Q$ proves. Suppose $x$ is arbitrary and $\text{Prf}_T(x, \gamma \rho_T^{-})$. We already know that $T \not \vdash \rho_T$, and so for every $k$, $Q \vdash \neg \text{Prf}_T(k, \gamma \rho_T^{-})$. Thus, for every $k$ it follows that $x \neq k$. In particular, we have (a) that $x \neq \pi$. We also have $\neg(x = 0 \lor x = 1 \lor \cdots \lor x = \overline{n} - 1)$ and so by $\neg \exists \overline{n} (2)$, (b) $\neg(x < \pi)$. By $\neg \exists \overline{n} (2)$, we have $\pi < x \land \text{Ref}_T(\pi, \gamma \rho_T^{-})$, and from that $\exists z (z < x \land \text{Ref}_T(z, \gamma \rho_T^{-}))$. Since $x$ was arbitrary we get

$$
\forall x (\text{Prf}_T(x, \gamma \rho_T^{-}) \rightarrow \exists z (z < x \land \text{Ref}_T(z, \gamma \rho_T^{-})))
$$

as required. 

\[\square\]

**Photo Credits**

**Bibliography**