

inp.1 Rosser's Theorem

inc:inp:ros:
sec 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” **derivability** predicate $\text{RProv}_T(y)$ instead of $\text{Prov}_T(y)$.

inc:inp:ros:
thm:rosser **Theorem inp.1.** *Let \mathbf{T} be any consistent, axiomatizable theory extending \mathbf{Q} . Then \mathbf{T} is not complete.*

Proof. Recall that $\text{Prov}_T(y)$ is defined as $\exists x \text{Prf}_T(x, y)$, where $\text{Prf}_T(x, y)$ represents the decidable relation which holds iff 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 $\text{not}(x)$ be the primitive recursive function which does the following: if x is the code of a formula φ , $\text{not}(x)$ is a code of $\neg\varphi$. Then $\text{Ref}_T(x, y)$ holds iff $\text{Prf}_T(x, \text{not}(y))$. Let $\text{Ref}_T(x, y)$ represent it. Then, if $\mathbf{T} \vdash \neg\varphi$ and δ is a corresponding **derivation**, $\mathbf{Q} \vdash \text{Ref}_T(\ulcorner\delta\urcorner, \ulcorner\varphi\urcorner)$. We define $\text{RProv}_T(y)$ as

$$\exists x (\text{Prf}_T(x, y) \wedge \forall z (z < x \rightarrow \neg \text{Ref}_T(z, y))).$$

Roughly, $\text{RProv}_T(y)$ says “there is a proof of y in \mathbf{T} , and there is no shorter refutation of y .” Assuming \mathbf{T} is consistent, $\text{RProv}_T(y)$ is true of the same numbers as $\text{Prov}_T(y)$; but from the point of view of *provability* in \mathbf{T} (and we now know that there is a difference between truth and provability!) the two have different properties. If \mathbf{T} is *inconsistent*, then the two do *not* hold of the same numbers! ($\text{RProv}_T(y)$ is often read as “ y is Rosser provable.” Since, as just discussed, Rosser provability is not some special kind of provability—in inconsistent theories, there are **sentences** that are provable but not Rosser provable—this may be confusing. To avoid the confusion, you could instead read it as “ y is shmovable.”)

By the fixed-point lemma, there is a formula $\rho_{\mathbf{T}}$ such that

$$\mathbf{Q} \vdash \rho_{\mathbf{T}} \leftrightarrow \neg \text{RProv}_T(\ulcorner\rho_{\mathbf{T}}\urcorner). \quad (1)$$

In contrast to the proof of ??, here we claim that if \mathbf{T} is consistent, \mathbf{T} doesn't **derive** $\rho_{\mathbf{T}}$, and \mathbf{T} also doesn't **derive** $\neg\rho_{\mathbf{T}}$. (In other words, we don't need the assumption of ω -consistency.)

First, let's show that $\mathbf{T} \not\vdash \rho_{\mathbf{T}}$. Suppose it did, so there is a **derivation** of $\rho_{\mathbf{T}}$ from T ; let n be its Gödel number. Then $\mathbf{Q} \vdash \text{Prf}_T(\bar{n}, \ulcorner\rho_{\mathbf{T}}\urcorner)$, since Prf_T represents Prf_T in \mathbf{Q} . Also, for each $k < n$, k is not the Gödel number of $\neg\rho_{\mathbf{T}}$, since \mathbf{T} is consistent. So for each $k < n$, $\mathbf{Q} \vdash \neg \text{Ref}_T(\bar{k}, \ulcorner\rho_{\mathbf{T}}\urcorner)$. By ??, $\mathbf{Q} \vdash \forall z (z < \bar{n} \rightarrow \neg \text{Ref}_T(z, \ulcorner\rho_{\mathbf{T}}\urcorner))$. Thus,

$$\mathbf{Q} \vdash \exists x (\text{Prf}_T(x, \ulcorner\rho_{\mathbf{T}}\urcorner) \wedge \forall z (z < x \rightarrow \neg \text{Ref}_T(z, \ulcorner\rho_{\mathbf{T}}\urcorner))),$$

but that's just $\text{RProv}_T(\ulcorner \rho_T \urcorner)$. By eq. (1), $\mathbf{Q} \vdash \neg \rho_T$. Since \mathbf{T} extends \mathbf{Q} , also $\mathbf{T} \vdash \neg \rho_T$. We've assumed that $\mathbf{T} \vdash \rho_T$, so \mathbf{T} would be inconsistent, contrary to the assumption of the theorem.

Now, let's show that $\mathbf{T} \not\vdash \neg \rho_T$. 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 Ref_T represents Ref_T in \mathbf{Q} , $\mathbf{Q} \vdash \text{Ref}_T(\bar{n}, \ulcorner \rho_T \urcorner)$. We'll again show that \mathbf{T} would then be inconsistent because it would also derive ρ_T . Since

$$\mathbf{Q} \vdash \rho_T \leftrightarrow \neg \text{RProv}_T(\ulcorner \rho_T \urcorner),$$

and since \mathbf{T} extends \mathbf{Q} , it suffices to show that

$$\mathbf{Q} \vdash \neg \text{RProv}_T(\ulcorner \rho_T \urcorner).$$

The sentence $\neg \text{RProv}_T(\ulcorner \rho_T \urcorner)$, i.e.,

$$\neg \exists x (\text{Prf}_T(x, \ulcorner \rho_T \urcorner) \wedge \forall z (z < x \rightarrow \neg \text{Ref}_T(z, \ulcorner \rho_T \urcorner)))$$

is logically equivalent to

$$\forall x (\text{Prf}_T(x, \ulcorner \rho_T \urcorner) \rightarrow \exists z (z < x \wedge \text{Ref}_T(z, \ulcorner \rho_T \urcorner)))$$

We argue informally using logic, making use of facts about what \mathbf{Q} derives. Suppose x is arbitrary and $\text{Prf}_T(x, \ulcorner \rho_T \urcorner)$. We already know that $\mathbf{T} \not\vdash \rho_T$, and so for every k , $\mathbf{Q} \vdash \neg \text{Prf}_T(\bar{k}, \ulcorner \rho_T \urcorner)$. Thus, for every k it follows that $x \neq \bar{k}$. In particular, we have (a) that $x \neq \bar{n}$. We also have $\neg(x = \bar{0} \vee x = \bar{1} \vee \dots \vee x = \overline{n-1})$ and so by ??, (b) $\neg(x < \bar{n})$. By ??, $\bar{n} < x$. Since $\mathbf{Q} \vdash \text{Ref}_T(\bar{n}, \ulcorner \rho_T \urcorner)$, we have $\bar{n} < x \wedge \text{Ref}_T(\bar{n}, \ulcorner \rho_T \urcorner)$, and from that $\exists z (z < x \wedge \text{Ref}_T(z, \ulcorner \rho_T \urcorner))$. Since x was arbitrary we get, as required, that

$$\forall x (\text{Prf}_T(x, \ulcorner \rho_T \urcorner) \rightarrow \exists z (z < x \wedge \text{Ref}_T(z, \ulcorner \rho_T \urcorner))). \quad \square$$

Photo Credits

Bibliography