## inp.1 Rosser's Theorem

inc:inp:ros:sec

Can we modify Gödel's proof to get a stronger result, replacing "$\omega$-consistent" with simply "consistent"? The answer is "yes," using a trick discovered by Rosser. Rosser's trick is to use a "modified" provability predicate $\mathsf{RProv}_T(y)$ instead of $\mathsf{Prov}_T(y)$.

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

$$\exists x\, (\mathsf{Prf}_T(x, y) \land \forall z\, (z < x \to \neg\mathsf{Ref}_T(z, y))).$$

Roughly, $\mathsf{RProv}_T(y)$ says "there is a proof of $y$ in $\mathbf{T}$, and there is no shorter refutation of $y$." (You might find it convenient to read $\mathsf{RProv}_T(y)$ as "$y$ is shmovable.") Assuming $\mathbf{T}$ is consistent, $\mathsf{RProv}_T(y)$ is true of the same numbers as $\mathsf{Prov}_T(y)$; 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 *in*consistent, then the two do *not* hold of the same numbers!)

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

inc:inp:ros:RT

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

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

First, let's show that $\mathbf{T} \nvdash \rho_T$. Suppose it did, so there is a derivation of $\rho_T$ from $T$; let $n$ be its Gödel number. Then $\mathbf{Q} \vdash \mathsf{Prf}_T(\overline{n}, \ulcorner\rho_T\urcorner)$, since $\mathsf{Prf}_T$ represents $\mathrm{Prf}_T$ in $\mathbf{Q}$. Also, for each $k < n$, $k$ is not the Gödel number of $\neg\rho_T$, since $\mathbf{T}$ is consistent. So for each $k < n$, $\mathbf{Q} \vdash \neg\mathsf{Ref}_T(\overline{k}, \ulcorner\rho_T\urcorner)$. By **??**(2), $\mathbf{Q} \vdash \forall z\, (z < \overline{n} \to \neg\mathsf{Ref}_T(z, \ulcorner\rho_T\urcorner))$. Thus,

$$\mathbf{Q} \vdash \exists x\, (\mathsf{Prf}_T(x, \ulcorner\rho_T\urcorner) \land \forall z\, (z < x \to \neg\mathsf{Ref}_T(z, \ulcorner\rho_T\urcorner))),$$

but that's just $\mathsf{RProv}_T(\ulcorner\rho_T\urcorner)$. 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 $\mathbf{T} \nvdash \neg\rho_T$. Again, suppose it did, and suppose $n$ is the Gödel number of a derivation of $\neg\rho_T$. Then $\mathrm{Ref}_T(n, {}^{\#}\rho_T{}^{\#})$ holds, and since $\mathsf{Ref}_T$ represents $\mathrm{Ref}_T$ in $\mathbf{Q}$, $\mathbf{Q} \vdash \mathsf{Ref}_T(\overline{n}, \ulcorner\rho_T\urcorner)$. We'll again show that $\mathbf{T}$ would then be inconsistent because it would also prove $\rho_T$. Since $\mathbf{Q} \vdash \rho_T \leftrightarrow \neg\mathsf{RProv}_T(\ulcorner\rho_T\urcorner)$, and since $\mathbf{T}$ extends $\mathbf{Q}$, it suffices to show that $\mathbf{Q} \vdash \neg\mathsf{RProv}_T(\ulcorner\rho_T\urcorner)$. The sentence $\neg\mathsf{RProv}_T(\ulcorner\rho_T\urcorner)$, i.e.,

$$\neg\exists x\,(\mathsf{Prf}_T(x, \ulcorner\rho_T\urcorner) \land \forall z\,(z < x \to \neg\mathsf{Ref}_T(z, \ulcorner\rho_T\urcorner)))$$

is logically equivalent to

$$\forall x\,(\mathsf{Prf}_T(x, \ulcorner\rho_T\urcorner) \to \exists z\,(z < x \land \mathsf{Ref}_T(z, \ulcorner\rho_T\urcorner)))$$

We argue informally using logic, making use of facts about what $\mathbf{Q}$ proves. Suppose $x$ is arbitrary and $\mathsf{Prf}_T(x, \ulcorner\rho_T\urcorner)$. We already know that $\mathbf{T} \nvdash \rho_T$, and so for every $k$, $\mathbf{Q} \vdash \neg\mathsf{Prf}_T(\overline{k}, \ulcorner\rho_T\urcorner)$. Thus, for every $k$ it follows that $x \neq \overline{k}$. In particular, we have (a) that $x \neq \overline{n}$. We also have $\neg(x = \overline{0} \lor x = \overline{1} \lor \dots \lor x = \overline{n-1})$ and so by **??**(2), (b) $\neg(x < \overline{n})$. By **??**, $\overline{n} < x$. Since $\mathbf{Q} \vdash \mathsf{Ref}_T(\overline{n}, \ulcorner\rho_T\urcorner)$, we have $\overline{n} < x \land \mathsf{Ref}_T(\overline{n}, \ulcorner\rho_T\urcorner)$, and from that $\exists z\,(z < x \land \mathsf{Ref}_T(z, \ulcorner\rho_T\urcorner))$. Since $x$ was arbitrary we get

$$\forall x\,(\mathsf{Prf}_T(x, \ulcorner\rho_T\urcorner) \to \exists z\,(z < x \land \mathsf{Ref}_T(z, \ulcorner\rho_T\urcorner)))$$

as required. □

## Photo Credits

## Bibliography