THE JOURNAL OF SYMBOLIC LOGIC
Volume 59, Number 4, December 1994

ON THE INCOMPLETENESS THEOREMS

HENRYK KOTLARSKI

Abstract. We give new proofs of both incompleteness theorems. We do not use the diagonalization lemma, but work with some quickly growing functions instead.

In the classical proofs of Gödel's incompleteness theorems one point, the diagonalization lemma, when used as a method of constructing an independent statement, is intuitively unclear (at least from the model-theoretic point of view). On the other hand, many results of this sort may be proved either by using diagonalization or by using some quickly growing functions. Therefore it seems to be of some interest to give proofs of both incompleteness theorems using quickly growing functions; such arguments are presented below. Some sort of diagonalization occurs in the proof when we are comparing two functions.

We have tried to make the paper as model-theoretic as possible. The reason is that (at least from the author's point of view) model-theoretic arguments are intuitively clearer than proof-theoretic ones.

We assume the reader to be familiar with arithmetization of syntax and with some model-theoretic constructions. Feferman [F], Hájek-Pudlák [HP], Kaye [K], Smoryński [Sm], or Shoenfield [Sh] contain all the necessary information. Smoryński's survey [Sm] was the main inspiration for us.

Acknowledgment. The author would like to thank several colleagues for discussions and e-mail correspondence, in particular Teresa Bigorajska, Richard Kaye, Roman Kossak, Jan Krajíček, and Jim Schmerl.

This paper is dedicated to the memory of Zygmunt Ratajczyk, who taught me a lot of things on the border of model theory and proof theory when we were much younger. He died in February 1994.

Let $\mathrm{Prov}_T(x, y)$ be the usual formula which expresses "x is a proof of the statement y from the axioms of T" and let $\mathrm{Pr}_T(y)$ be $\exists x\ \mathrm{Prov}_T(x, y)$. Let $\mathrm{Con}_{\mathrm{PA}}$ be the statement which expresses "PA is consistent". Thus, $\mathrm{Con}_{\mathrm{PA}}$ is $\neg \mathrm{Pr}_{\mathrm{PA}}(0 = 1)$. Let us also denote Tr_0 the usual universal formula for Δ_0-formulas. Thus we have:

$$\text{(1)} \qquad \text{for every } \varphi \in \Delta_0 \quad \mathrm{PA} \vdash \forall b[\varphi(b) \equiv \mathrm{Tr}_0(\varphi(S^b 0))].$$

Here by $S^b 0$ we denote the bth numeral, i.e.

$$S^b 0 = \overbrace{S \cdots S}^{b\text{times}} 0.$$

Received December 14, 1993; revised May 3, 1994.

Research supported in part by The State Committee for Scientific Research (Poland), KBN, grant number 2 2029 92 03.

0022-4812/94/5904-0020/$01.60

The proof of (1), by induction on φ, is well known. It is effective in the sense that as a matter of fact it gives a primitive recursive function which assigns to every $\varphi \in \Delta_0$ the (Gödel number of the) proof of the desired statement $\forall b[\varphi(b) \equiv \mathrm{Tr}_0(\varphi(S^b 0))]$. It follows that (1) itself is provable in PA. Thus we have

LEMMA 1. *PA proves the following statement*:

$$\forall \varphi \in \Delta_0 \quad \mathrm{Pr}_{\mathrm{PA}}(\forall b[\varphi(b) \equiv \mathrm{Tr}_0(\varphi(S^b 0))]). \qquad \square$$

The following remark may help the reader's intuition. Below we shall work in nonstandard models of PA. Lemma 1 ensures that if $\mathcal{M} \models PA$, then for every object in $\mathcal{M}$ which $\mathcal{M}$ is a (Gödel number of a) Δ_0 formula, there exists an object in $\mathcal{M}$ which $\mathcal{M}$ thinks is a Gödel number of a proof of the statement $\forall b[\varphi(b) \equiv \mathrm{Tr}_0(\varphi(S^b 0))]$. None of these objects need be standard, they are just elements of $\mathcal{M}$. Similar remarks apply to all statements that we claim are provable in PA.

We shall need some more observations about the connection between truth and provability of Δ_0-formulas in PA. The next two lemmas are well known.

LEMMA 2. *PA proves the following statement*:

$$\forall \varphi \in \Delta_0 \ \forall b[\mathrm{Tr}_0(\varphi(S^b 0)) \Rightarrow \mathrm{Pr}_{\mathrm{PA}}(\varphi(S^b 0))].$$

IDEA OF THE PROOF. By induction on φ one constructs a proof of the desired statement

$$\forall b[\mathrm{Tr}_0(\varphi(S^b 0)) \Rightarrow \mathrm{Pr}_{\mathrm{PA}}(\varphi(S^b 0))]$$

primitive recursively in φ. Simply the assumption $\mathrm{Tr}_0(\varphi(S^b 0))$ gives (essentially) a computation of the logical value of $\varphi(S^b 0)$ and hence a proof of this statement in sentential calculus. $\square$

LEMMA 3. *PA proves the following statement*:

$$\mathrm{Con}_{\mathrm{PA}} \Rightarrow \forall \varphi \in \Delta_0 \ \forall b[\mathrm{Pr}_{\mathrm{PA}}(\exists w \leq S^b 0 \ \ \mathrm{Tr}_0(\varphi(S^b 0))) \Rightarrow \exists w \leq b \ \ \mathrm{Tr}_0(\varphi(S^w 0))].$$

PROOF. Assume $\mathrm{Pr}_{\mathrm{PA}}(\exists w \leq S^b 0 \ \mathrm{Tr}_0(\varphi(S^w 0)))$. If $\forall w \leq b \ \mathrm{Tr}_0(\neg\varphi(S^w 0))$, then, by Lemma 2, we infer that $\mathrm{Pr}_{\mathrm{PA}}(\forall w \leq b \ \neg\varphi(S^w 0))$, so $\neg\,\mathrm{Con}_{\mathrm{PA}}$. $\square$

DEFINITION (in PA).

$$F(a) = \min b \colon \forall \varphi, u \leq a\{[\varphi \in \Delta_0 \wedge \exists w \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0))] \\ \Rightarrow \exists w < b \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0))\}.$$

Thus F is the natural function which dominates all Δ_0 functions. It is easy to prove in PA by induction the statement $\forall a \ \exists b \ b = F(a)$. The definition of F is not Δ_0 (because of the quantifier $\exists w$). As we shall see, this is the heart of the matter in the proofs of the incompleteness theorems.

LEMMA 4. *There exists a natural number a_0 such that PA proves*

$$\mathrm{Con}_{\mathrm{PA}} \ \Rightarrow \ \forall b \ \neg\,\mathrm{Pr}_{\mathrm{PA}}(F(S^{a_0} 0) \leq S^b 0).$$

PROOF. We define the following function:

$$G(a) = \min\langle x, b, z, d\rangle \colon d \text{ witnesses that } z \text{ is a substitution} \\ \text{of the form } F(S^a 0) \leq S^b 0 \text{ and } \mathrm{Prov}_{\mathrm{PA}}(x, z).$$

Of course, $\langle x, b, z, d\rangle$ denotes the tuple whose items are x, b, z, d. Observe that the definition of G is Δ_0 as written. We put $a_0 =$ the Gödel number of the formula $v_2 = 1 + G(v_1)$ and shall verify that this number a_0 satisfies our demand. But first we need a preparatory remark.

Let $\mathscr{M} \models \mathrm{PA} + \mathrm{Con}_{\mathrm{PA}}$. Work inside $\mathscr{M}$. Fix $a \in \mathscr{M}$ and let $G(a) = \langle x, b, z, d\rangle$. Let $\varphi, u \leq a$ be given with $\varphi \in \Delta_0$. Assume

$$\exists w \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0)).$$

Then $\exists w \ \ \mathrm{Pr}_{\mathrm{PA}}(\varphi(S^u 0, S^w 0))$ by Lemma 2, and hence $\mathrm{Pr}_{\mathrm{PA}}(\exists w \varphi(S^u 0, w))$ by the $\exists$-introduction rule (for which the appropriate derivability condition holds). By Lemma 1 we infer that

$$\mathrm{Pr}_{\mathrm{PA}}(\exists w \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0)))$$

and, by definition of F,

$$\mathrm{Pr}_{\mathrm{PA}}(\exists w < S^b 0 \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0))).$$

It follows that $\mathscr{M} \models \exists w \leq b \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0))$ by Lemma 3. Summing up, we see that $\mathscr{M}$ satisfies

$$\begin{aligned} &\text{if } G(a) \text{ exists then} \\ &\quad \forall \varphi, u \leq a\{[\varphi \in \Delta_0 \wedge \exists w \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0))] \\ &\qquad\qquad \Rightarrow \exists w \leq G(a) \ \ \mathrm{Tr}_0(\varphi(S^u 0, S^w 0))\}, \end{aligned}$$

because $b \leq G(a)$. We let $\varphi = a_0$ and the parameter $u = a_0$. Thus we infer that

$$\exists w \leq G(a_0) \ \ \mathrm{Tr}_0(S^w 0 = 1 + G(a_0)),$$

i.e. $1 + G(a_0) \leq G(a_0)$, contradiction. Thus $G(a_0)$ cannot exist. □

Observe that a_0 played two roles in the proof of Lemma 4: it was used as a (Gödel number of a) formula and a parameter. The diagonalization procedure occurred in the proof of Lemma 4; indeed, the function F was used in $\mathscr{M}$ and inside $\mathrm{Prov}_{\mathrm{PA}}$—the heart of the matter was just comparison of the rate of growth of these two versions of F.

Before going further let us describe the idea of the so-called arithmetized completeness theorem. We follow Smoryński's presentation [Sm], with some minor changes. Let Tr_2 denote the usual universal formula for Σ_2 formulas. Let $\mathrm{Compl}(C)$ denote the formula which expresses "C is the Gödel number of some Σ_2-formula which describes a complete and consistent extension of PA". Thus $\mathrm{Compl}(C)$ is

$$\begin{aligned} &C \in \Sigma_2 \wedge \forall x[\mathrm{Tr}_2(C; x) \Rightarrow \mathrm{Sent}(x)] \\ &\qquad \wedge \forall \varphi\{\mathrm{Sent}(\varphi) \Rightarrow [\mathrm{Tr}_2(C; \varphi) \vee \mathrm{Tr}_2(C; \neg\varphi)]\} \\ &\qquad \wedge \forall \langle \varphi_0, \dots, \varphi_{r-1}\rangle \left\{[\forall i < r \ \ \mathrm{Tr}_2(C; \varphi_i)] \Rightarrow \neg \mathrm{Pr}_{\mathrm{PA}}\left(\neg \bigwedge_{i<r} \varphi_i\right)\right\}. \end{aligned}$$

Once again we want to point out that if we are given a model $\mathscr{M}$ for PA and $C \in \mathscr{M}$ satisfying Compl then C need not be a standard object. This is just an element of $\mathscr{M}$ which $\mathscr{M}$ thinks is the Gödel number of a Σ_2-formula. Observe that $\mathrm{Compl}(\cdot)$ is Π_3.

The following fact is known as the Hilbert-Bernays arithmetized completeness theorem.

LEMMA 5. *PA proves* $\mathrm{Con}_{\mathrm{PA}} \equiv \exists C\ \mathrm{Compl}(C)$.

PROOF. See Smoryński [Sm]. □

Suppose we are given $\mathscr{M} \models \mathrm{PA}$, and let $\mathscr{M} \models \mathrm{Compl}(C)$, i.e. C is a completion of PA in $\mathscr{M}$. These data determine a new model. It is constructed as follows. Let $\mathrm{Tm}^{\mathscr{M}}$ denote the set of all (Skolem) constant terms in the sense of $\mathscr{M}$. Divide it by the equivalence relation $t_1 \sim t_2 \equiv \mathscr{M} \models \mathrm{Tr}_2(C; t_1 = t_2)$. Clearly this is an equivalence relation, and the following definition of addition makes sense:

$$t_1^{\sim} + t_2^{\sim} = t_3^{\sim} \quad \equiv \quad \mathscr{M} \models \mathrm{Tr}_2(C; t_1 + t_2 = t_3).$$

We treat other atomic symbols in the language of PA similarly. We denote by $\mathrm{ACT}(\mathscr{M}; C)$ the model constructed above. ACT stands for the arithmetized completeness theorem.

LEMMA 6. *Let* $\mathscr{M}$ *be a model of PA and let* C *be a completion in* $\mathscr{M}$. *Then for every formula* $A(v_0, \ldots, v_{r-1})$ *and* r*-tuple* $t_0, \ldots, t_{r-1}$ *we have*

$$\mathrm{ACT}(\mathscr{M}; C) \models A(t_0^{\sim}, \ldots, t_{r-1}^{\sim}) \quad \text{iff} \quad \mathscr{M} \models \mathrm{Tr}_2(C; A(t_0, \ldots, t_{r-1})).$$

COMMENTS ON THE PROOF. This is a standard Henkin-like argument. Let us give one minor observation. The function which associates to every formula its Skolem term (given by the scheme of minimum) is primitive recursive, so we can work with it freely inside $\mathscr{M}$. It follows that C has in $\mathscr{M}$ the properties of a complete Skolemized theory, so the usual argument works smoothly. □

Let $\mathscr{M} \models \mathrm{PA}$ and let C be a completion of PA in $\mathscr{M}$. It turns out that there exists a natural embedding of $\mathscr{M}$ onto an initial segment of $\mathrm{ACT}(\mathscr{M}; C)$. It is defined as follows: we map $b \in \mathscr{M}$ to the equivalence class of the numeral $S^b 0$ in $\mathrm{ACT}(\mathscr{M}; C)$. Let j denote this embedding. Thus we have

LEMMA 7. *If* $A(v_0, \ldots, v_{r-1})$ *is a* Δ_0 *formula and* $b_0, \ldots, b_{r-1} \in \mathscr{M}$, *then*

$$\mathscr{M} \models A(b_0, \ldots, b_{r-1}) \quad \text{iff} \quad \mathrm{ACT}(\mathscr{M}; C) \models A(j(b_0), \ldots, j(b_{r-1})).$$

The same absoluteness holds for Δ_1 *formulas, etc.* □

Let a be a natural number with the property stated in Lemma 4. Let $\mathscr{M}$ be a model for $\mathrm{PA} + \mathrm{Con}_{\mathrm{PA}}$. Then

$$\mathscr{M} \models \forall b\ \neg \mathrm{Pr}_{\mathrm{PA}}(F(S^a 0) \leq S^b 0)$$

and by an inessential variant of Lemma 5 there exists a completion C in $\mathscr{M}$ such that

$$\mathscr{M} \models \forall b\ \ \mathrm{Tr}_2(C; F(S^a 0) > S^b 0).$$

It is convenient to think of this phenomenon as follows: $F(a)$ in the sense of

$\mathrm{ACT}(\mathscr{M}; C)$ is not $\mathscr{M}, C$-standard. By the definition of F this means for some $\varphi, u \leq a$, $\varphi \in \Delta_0$,

$$\forall w \in \mathscr{M} \;\; \mathrm{ACT}(\mathscr{M}; C) \models \neg\varphi(u, w)$$

but

$$\mathrm{ACT}(\mathscr{M}; C) \models \exists x \;\; \varphi(u, x).$$

But $\varphi \in \Delta_0$, so by the absoluteness lemma (i.e. Lemma 7) $\mathscr{M} \models \neg\varphi(u, w)$, and hence

$$\mathscr{M} \models \forall w \;\; \mathrm{Pr}_{\mathrm{PA}}(\neg\varphi(S^u 0, S^w 0)).$$

Thus, if we assume that $\mathscr{M} \models$ "PA is ω-consistent (with respect to Δ_0 formulas)", then

$$\mathscr{M} \models \neg \mathrm{Pr}_{\mathrm{PA}}(\exists x \;\; \varphi(S^u 0, x)).$$

On the other hand, we know that $\mathscr{M} \models \neg \mathrm{Pr}_{\mathrm{PA}}(\neg\exists x \, \varphi(S^u 0, x))$, because the statement $\exists x \, \varphi(S^u 0, x))$ is in C. Let us sum up.

THEOREM 8 (The First Incompleteness Theorem). *PA proves*

$$\textit{if } \forall\varphi, u\{[\varphi \in \Delta_0 \wedge \forall w \;\; \mathrm{Pr}_{\mathrm{PA}}(\neg\varphi(S^u 0, S^w 0))] \Rightarrow \neg \mathrm{Pr}_{\mathrm{PA}}(\exists x \;\; \varphi(S^u 0, x))\}$$
$$\textit{then } \exists\varphi \in \Delta_0 \; \exists u[\neg \mathrm{Pr}_{\mathrm{PA}}(\exists x \;\; \varphi(S^u 0, x)) \wedge \neg \mathrm{Pr}_{\mathrm{PA}}(\neg\exists x \;\; \varphi(S^u 0, x))].$$

In particular, if we apply Theorem 8 *inside the standard model then we obtain a* Σ_1 *independent statement.* □

In order to derive the second incompleteness theorem from Lemma 4 we need some minor additional work. The reason is that in the proof of Theorem 8 we ensured that $F(a)$ in the sense of $\mathrm{ACT}(\mathscr{M}; C)$ is not $\mathscr{M}$-standard. But this could be caused by another formula than "there exists a proof $0 = 1$ in PA"—it could be caused, say, by the formula "there exists a proof of contradiction of ZF". Let us show how to overcome this difficulty.

THEOREM 9 (The Second Incompleteness Theorem). *PA does not prove* $\mathrm{Con}_{\mathrm{PA}}$.

PROOF. Assume the contrary: PA proves its own consistency. Let a be a natural number with the property stated in Lemma 4. Enumerate

$$\varphi_0(S^{u_0} 0, x), \ldots, \varphi_{r-1}(S^{u_{r-1}} 0, x)$$

all substitutions with $\varphi, u \leq a$ and $\varphi \in \Delta_0$. Clearly there are at most $(a+1)^2$ such substitutions. We iterate the construction of a new model by the arithmetized completeness theorem. So let $\mathscr{M} = \mathscr{M}_0$ be a model of PA. Consider the first substitution $\varphi_0(S^{u_0} 0, x)$. If $\mathscr{M} \models \mathrm{Pr}_{\mathrm{PA}}(\forall x \;\; \neg\varphi_0(S^{u_0} 0, x))$ then let $\mathscr{M}_1 = \mathscr{M}$. Otherwise

$$\mathscr{M} \models \neg \mathrm{Pr}_{\mathrm{PA}}(\neg\exists x \;\; \varphi_0(S^{u_0} 0, x)),$$

and hence there exists a completion C in $\mathscr{M}$ such that $\mathrm{ACT}(\mathscr{M}; C) \models \exists x \; \varphi_0(S^{u_0} 0, x)$. Thus there exists $\mathscr{M}_1 \supseteq_e \mathscr{M}$ such that

$$\text{either } \mathscr{M}_1 \models \mathrm{Pr}_{\mathrm{PA}}(\neg\exists x \;\; \varphi_0(S^{u_0} 0, x))$$
$$\text{or } \mathscr{M}_1 \models \exists w \;\; \mathrm{Pr}_{\mathrm{PA}}(\varphi_0(S^{u_0} 0, S^w 0)).$$

(Here we use the fact that φ is Δ_0, hence its truth in $\mathscr{M}_1$ implies it provability.) We iterate this construction (consider now $\mathscr{M}_1$ and $\varphi_1(S^{u_1}0, x)$, etc.). This induction either breaks before $r-1$ steps, and hence we get a model for $\mathrm{PA} + \neg \mathrm{Con}_{\mathrm{PA}}$, or the final model $\mathscr{M}_{r-1}$ satisfies $\neg \mathrm{Con}_{\mathrm{PA}}$ by Lemma 4. □

Let us remark that the so-called formalized second incompleteness theorem is also a consequence of the construction presented above. In order to be a bit more precise, we claim that PA proves the following statement:

$$\mathrm{Con}_{\mathrm{PA}} \ \Rightarrow \ \mathrm{Con}_{\mathrm{PA} + \neg \mathrm{Con}_{\mathrm{PA}}}.$$

In order to see why it is so, we shall show that every model $\mathscr{M}_j$ from the proof of Theorem 9 either is equal to $\mathscr{M}$ or has an elementary submodel of the form $\mathrm{ACT}(\mathscr{M}; E)$ for some completion E in $\mathscr{M}$. Granted this, we see that $\mathrm{PA} + \neg \mathrm{Con}_{\mathrm{PA}}$ is contained in some completion E in $\mathscr{M}$, so is consistent. Moreover this consistency holds in every $\mathscr{M}$ which is a model for $\mathrm{PA} + \mathrm{Con}_{\mathrm{PA}}$, so is provable in this theory. In order to verify the above property of the chain $\mathscr{M}_j$ of models considered, it suffices (by induction) to check the following.

Observation 10. Let $\mathscr{M} \models PA$, let C be a completion in $\mathscr{M}$, and let D be a completion in $\mathrm{ACT}(\mathscr{M}; C)$. Under these assumptions there exists a completion E in $\mathscr{M}$ so that

$$\mathrm{ACT}(\mathscr{M}; E) \prec \mathrm{ACT}(\mathrm{ACT}(\mathscr{M}; C); D).$$

IDEA OF THE PROOF. Let $\mathscr{M}$, C, and D satisfy the assumption. Let j denote the embedding determined by $\mathscr{M}$ and C in the manner described above. We put

$$\psi \in E \quad \text{iff} \quad \exists z[z = j(\psi) \wedge \mathrm{Tr}_2(C; \mathrm{Tr}_2(D; z))].$$

It requires some minor work to check that E satisfies the conclusion; we leave it to the reader, but just mention a fine point. In order to check that E is Σ_2 one replaces the definition of j by an inductive one. □

It is not clear at the moment whether the method presented above will also give a new proof of Rosser's theorem.

REFERENCES

[F] S. FEFERMAN, *Arithmetization of metamathematics in a general setting*, ***Fundamenta Mathematicae***, vol. 49 (1960), pp. 35–92.

[HP] P. HÁJEK and P. PUDLÁK, *Metamathematics of first order arithmetic*, ***Perspectives in Mathematical Logic***, Springer Verlag, Berlin, 1993.

[K] R. KAYE, *Models of Peano arithmetic*, ***Oxford Logic Guides***, vol. 15, Oxford University Press, Oxford, 1991.

[Sh] J. R. SHOENFIELD, ***Mathematical logic***, Addison-Wesley, Reading, Massachusetts, 1967.

[Sm] C. SMORYŃSKI, *The incompleteness theorems*, ***Handbook of mathematical logic*** (J. Barwise, editor), North-Holland, Amsterdam, 1977, pp. 821–865.

KATEDRA MATEMATYKI WSRP W SIEDLCACH
UL. ORLICZ-DRESZERA 19/21
08-110 SIEDLCE, POLAND

E-mail: klarski@plearn.bitnet