Annals of Pure and Applied Logic 169 (2018) 971–980

Contents lists available at ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

On constructivity and the Rosser property: a closer look at some Gödelean proofs [☆]

Saeed Salehi [a,b,*], Payam Seraji [c]

[a] *Research Institute for Fundamental Sciences, University of Tabriz, 29 Bahman Boulevard, P.O.Box 51666-16471, Tabriz, Iran*
[b] *School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P.O.Box 19395-5746, Tehran, Iran*
[c] *Department of Mathematical Sciences, University of Tabriz, 29 Bahman Boulevard, P.O.Box 51666-16471, Tabriz, Iran*

ARTICLE INFO

Article history:
Received 1 May 2017
Received in revised form 15 January 2018
Accepted 23 April 2018
Available online 29 May 2018

MSC:
03F40
03F30
03D32
68Q30

Keywords:
Constructive proof
Rosser Property
Incompleteness
Kleene's proof
Chaitin's proof
Boolos' proof

ABSTRACT

The proofs of Kleene, Chaitin and Boolos for Gödel's First Incompleteness Theorem are studied from the perspectives of constructivity and the Rosser property. A proof of the incompleteness theorem has the Rosser property when the independence of the true but unprovable sentence can be shown by assuming only the (simple) consistency of the theory. It is known that Gödel's own proof for his incompleteness theorem does not have the Rosser property, and we show that neither do Kleene's or Boolos' proofs. However, we show that a variant of Chaitin's proof can have the Rosser property. The proofs of Gödel, Rosser and Kleene are constructive in the sense that they explicitly construct, by algorithmic ways, the independent sentence(s) from the theory. We show that the proofs of Chaitin and Boolos are not constructive, and they prove only the mere existence of the independent sentences.

[☆] This (except the second section on Kleene's proof) is a part of the Ph.D. thesis of the second author written at the University of Tabriz, IRAN, under the supervision of the first author who is partially supported by grant № 94030033 from the Institute for Research in Fundamental Sciences IPM.

[*] Corresponding author at: Research Institute for Fundamental Sciences, University of Tabriz, 29 Bahman Boulevard, P.O.Box 51666-16471, Tabriz, Iran.

E-mail addresses: salehipour@tabrizu.ac.ir, saeedsalehi@ipm.ir (S. Salehi), payam.seraji54@gmail.com (P. Seraji).
URL: http://www.saeedsalehi.ir (S. Salehi).

https://doi.org/10.1016/j.apal.2018.04.009

1. Introduction

A constructive proof provides an algorithm for constructing the claimed object; a non-constructive proof does not show the existence of that object algorithmically, even if sometimes an effective procedure might be hidden inside the details. A proof then *is proved to be* (*essentially*) *non-constructive* when one can show that there is no algorithm (computable function) which, given the assumptions (coded as input), produces the claimed object whose existence is demonstrated in the proof. Below, we will see one example of a (seemingly) non-constructive proof (namely, the proof of Kleene [12] for Gödel's first incompleteness theorem, stated below) which can be made constructive by unpacking some details; we will also see a couple of proofs (namely, the proofs of Boolos [2] and Chaitin [4] for Gödel's first incompleteness theorem) that are shown to be non-constructive, by proving the non-existence of any algorithm for computing the claimed object (namely, the true but unprovable sentence).

The (First) *Incompleteness Theorem* (of Gödel [6]) states that for a sufficiently strong RE theory T there exists a sentence $\boldsymbol{\psi}_T$ in the language of T such that

1. the sentence $\boldsymbol{\psi}_T$ is true (in the standard model of natural numbers);
2. if T is consistent then $T \nvdash \boldsymbol{\psi}_T$;
3. if T is ω-consistent then $T \nvdash \neg\boldsymbol{\psi}_T$.

By *a proof of the incompleteness theorem* we mean a demonstration of the existence of such a sentence ($\boldsymbol{\psi}_T$) for any given consistent and RE theory T that is sufficiently strong (to be made precise later). Such a proof witnesses *the Rosser property* ([17]) when the condition of ω-consistency can be replaced with (simple) consistency; that is to say that the condition 3 above can be replaced with the following condition

3′. if T is consistent then $T \nvdash \neg\boldsymbol{\psi}_T$.

Gödel's original proof [6] for his incompleteness theorem is constructive, i.e., given a (finite) description of a consistent RE theory (e.g. an input-free program which outputs the set of all the axioms of the theory) the proof exhibits, in an algorithmic way, a sentence which is true (in the standard model of natural numbers $\mathbb{N}$) but unprovable in the theory. For the independence of this sentence from the theory (i.e., the unprovability of its negation in the theory) Gödel also assumes the theory to be ω-consistent; so if the theory is ω-consistent, then that (true) sentence is independent from the theory (see e.g. [22,21]). It turned out later that the simple consistency of the theory does not suffice for the independence of the Gödel sentence (from the theory) and the optimal condition (which is much weaker than ω-consistency) is *the consistency of the theory with its own consistency statement* ([8, Theorems 35,36]). Rosser's proof [17] for Gödel's first incompleteness theorem assumes only the simple consistency of the (RE) theory and constructs (algorithmically) an independent (and true) sentence. So, one can say that Gödel's proof does not have the Rosser property. Here, we will see that while a variant of the proof of Chaitin has the Rosser property (i.e., the independence of Chaitin's sentence from the theory can be proved by assuming only the simple consistency of the theory), the proof of Boolos does not have the Rosser property (and the optimal condition for the independence of a Boolos sentence is the consistency of the theory with its own consistency statement).

2. The proof of Kleene for Gödel's incompleteness theorem

A very cute proof for Gödel's incompleteness theorem is that of Kleene (see e.g. [12,21]) which deserves more recognition.

Notation 2.1 *(Computability).* Let $\varphi_0, \varphi_1, \varphi_2, \cdots$ be a list of all unary computable (partial recursive) functions (in a way that $\varphi_i(j)$, if it exists, can be computed from i and j). A recursively enumerable set (RE for short) is the domain of φ_i, for some $i \in \mathbb{N}$, which is denoted by $\mathcal{W}_i$. The notation $\varphi_i(j)\!\uparrow$ means that the function φ_i is not defined at j, or $j \notin \mathcal{W}_i$; and $\varphi_i(j)\!\downarrow$ means that φ_i is defined at j or $j \in \mathcal{W}_i$. Needless to say, $\varphi_i(j) = k$ means that φ_i is defined at j and is equal to k. $\circledast$

Robinson's Arithmetic is denoted by Q (see [24] or [21]). In all the results of this paper, the theory Q can be replaced with a (much) weaker theory called R (see [24]). The theory Q is finitely axiomatizable, while R is not.

Theorem 2.2 *(Kleene's Theorem). For a given consistent and* RE *theory T that contains Q there exists some $t \in \mathbb{N}$ such that $\varphi_t(t)\!\uparrow$ but $T \nvdash$ "$\varphi_t(t)\!\uparrow$".*

Non-Constructive Proof. Let $\mathcal{K}_T = \{n \in \mathbb{N} \mid T \vdash \text{“}\varphi_n(n)\!\uparrow\text{”}\}$; then we have $\mathcal{K}_T \subseteq K = \{n \in \mathbb{N} \mid \varphi_n(n)\!\uparrow\}$ since if $T \vdash$ "$\varphi_n(n)\!\uparrow$" but $\varphi_n(n)\!\downarrow$ then the true Σ_1-sentence "$\varphi_n(n)\!\downarrow$" is provable in (the Σ_1-complete theory) Q ($\subseteq T$) contradicting the consistency of T. Now, since T is RE then so is $\mathcal{K}_T$, while K is not an RE set because for any n we have $n \in K \iff n \notin \mathcal{W}_n$ and so $n \in K \triangle \mathcal{W}_n$, thus $K \neq \mathcal{W}_n$ for all n. So, $\mathcal{K}_T \subsetneqq K$; therefore, there must exist some $t \in K - \mathcal{K}_T$. For this t we have $\varphi_t(t)\!\uparrow$ but $T \nvdash$ "$\varphi_t(t)\!\uparrow$". $\boxtimes$

Of course if T is sound (i.e., $\mathbb{N} \models T$) or even Σ_1-sound (i.e., if $\sigma \in \Sigma_1$ and $T \vdash \sigma$ then $\mathbb{N} \models \sigma$, cf. [8]) then also $T \nvdash$ "$\varphi_t(t)\!\downarrow$", i.e., the sentence "$\varphi_t(t)\!\uparrow$" is (true and) independent from T. Let us note that the above proof did not explicitly specify $t \in \mathbb{N}$.

Constructive Proof. Since $\mathcal{K}_T = \{n \in \mathbb{N} \mid T \vdash \text{“}\varphi_n(n)\!\uparrow\text{”}\}$ is RE then $\mathcal{K}_T = \mathcal{W}_{\mathtt{t}}$ for some $\mathtt{t} \in \mathbb{N}$ which can be algorithmically computed from a description of the RE theory T. Now we show the truth of "$\varphi_{\mathtt{t}}(\mathtt{t})\!\uparrow$" as follows:

$$
\begin{aligned}
\varphi_{\mathtt{t}}(\mathtt{t})\!\downarrow \quad &\Longrightarrow \quad T \vdash \text{“}\varphi_{\mathtt{t}}(\mathtt{t})\!\downarrow\text{”} && \text{(by the } \Sigma_1\text{-completeness of } \mathsf{Q} \subseteq T) \\
&\Longrightarrow \quad T \nvdash \text{“}\varphi_{\mathtt{t}}(\mathtt{t})\!\uparrow\text{”} && \text{(by the consistency of } T) \\
&\Longrightarrow \quad \mathtt{t} \notin \mathcal{K}_T && \text{(by the definition of } \mathcal{K}_T) \\
&\Longrightarrow \quad \mathtt{t} \notin \mathcal{W}_{\mathtt{t}} && \text{(by } \mathcal{K}_T = \mathcal{W}_{\mathtt{t}}) \\
&\Longrightarrow \quad \varphi_{\mathtt{t}}(\mathtt{t})\!\uparrow && \text{(by the definition of } \mathcal{W}_{\mathtt{t}})
\end{aligned}
$$

Thus, $\mathtt{t} \notin \mathcal{W}_{\mathtt{t}}$ and so $\mathtt{t} \notin \mathcal{K}_T$ whence $T \nvdash$ "$\varphi_{\mathtt{t}}(\mathtt{t})\!\uparrow$". $\boxtimes$

Indeed, for any RE and consistent theory $T(\supseteq \mathsf{Q})$ and *any* t with $\mathcal{W}_t = \mathcal{K}_T$ we have (by the above proof) that $\varphi_t(t)\!\uparrow$ and $T \nvdash$ "$\varphi_t(t)\!\uparrow$". Below we show that Kleene's (constructive) proof does not have the Rosser property.

Theorem 2.3 *(Kleene's Proof is not Rosserian). For any given consistent and* RE *theory $T \supseteq \mathsf{Q}$ there exists an* RE *and consistent theory $U \supseteq T$ such that $U \vdash$ "$\varphi_u(u)\!\downarrow$" for some $u \in \mathbb{N}$ which satisfies $\mathcal{W}_u = \{n \in \mathbb{N} \mid U \vdash \text{“}\varphi_n(n)\!\uparrow\text{”}\}$ (and $\varphi_u(u)\!\uparrow$).*

Proof. There exists a computable (and total) function $\hbar$ such that for any sentence ψ we have

$$\mathcal{W}_{\hbar(\psi)} = \{n \in \mathbb{N} \mid T + \psi \vdash \text{“}\varphi_n(n)\!\uparrow\text{”}\}.$$

By the Diagonal Lemma there exists a sentence λ such that $\mathsf{Q} \vdash \lambda \leftrightarrow$ "$\varphi_{\hbar(\lambda)}\big(\hbar(\lambda)\big)\!\downarrow$". Clearly, for the theory $U = T + \lambda$ and $u = \hbar(\lambda)$ we have $U \vdash$ "$\varphi_u(u)\!\downarrow$" and $\mathcal{W}_u = \{n \in \mathbb{N} \mid U \vdash \text{“}\varphi_n(n)\!\uparrow\text{”}\}$. It remains to show that U is consistent: Otherwise, $T \vdash \neg\lambda$ and so $T \vdash$ "$\varphi_u(u)\uparrow$" which implies that $T+\lambda \vdash$ "$\varphi_u(u)\uparrow$" whence $u \in \mathcal{W}_{\hbar(\lambda)} = \mathcal{W}_u$. On the other hand $T \vdash$ "$\varphi_u(u)\uparrow$" implies that $\varphi_u(u)\uparrow$ holds (since otherwise $\varphi_u(u)\downarrow$ by the Σ_1-completeness would imply $T \vdash$ "$\varphi_u(u)\downarrow$" contradicting the consistency of T) and so $u \notin \mathcal{W}_u$; a contradiction. ⊠

Summing up, for any consistent and RE extension T of Q we have $\varphi_t(t)\uparrow$ and $T \nvdash$ "$\varphi_t(t)\uparrow$" for any t which satisfies $\mathcal{W}_t = \{n \in \mathbb{N} \mid T \vdash$ "$\varphi_n(n)\uparrow$"$\}$. Moreover, if T is Σ_1-sound then $\varphi_t(t)\uparrow$ is independent from T (i.e., we also have $T \nvdash$ "$\varphi_t(t)\downarrow$"). However, if the theory T is not Σ_1-sound then for some e with $\mathcal{W}_e = \{n \in \mathbb{N} \mid T \vdash$ "$\varphi_n(n)\uparrow$"$\}$ the sentence $\varphi_e(e)\uparrow$ might not be independent from T (and its negation could be provable in T, that is $T \vdash$ "$\varphi_e(e)\downarrow$").

Albert Visser pointed out that for any RE and consistent theory T which is sufficiently strong (see e.g. the explanations before Theorem 4.6 below) there exists some ϑ with $\mathcal{W}_\vartheta = \{n \in \mathbb{N} \mid T \vdash$ "$\varphi_n(n)\uparrow$"$\}$ such that (beside $\varphi_\vartheta(\vartheta)\uparrow$ and $T \nvdash$ "$\varphi_\vartheta(\vartheta)\uparrow$" we also have) $T \nvdash$ "$\varphi_\vartheta(\vartheta)\downarrow$", or in the other words the sentence $\varphi_\vartheta(\vartheta)\uparrow$ is independent from T; moreover ϑ can be algorithmically computed from a given description of the RE theory T. The proof of this Rosserian version of Kleene's proof is rather involved and will appear in a future paper. Let us note that a Rosserian version of this beautiful theorem of Kleene appeared in [13] (see also [14]) where Kleene calls it "a symmetric form" of Gödel's (incompleteness) theorem (also see [19] for a modern treatment).

3. The proof of Chaitin for Gödel's incompleteness theorem

There are various versions of Chaitin's proof for the incompleteness theorem [4], which is sometimes called "Chaitin's incompleteness theorem"; this proof appears in e.g. [5,16,1,23]. We consider the version presented in [1].

Definition 3.1 *(Kolmogorov–Chaitin Complexity).* For any natural number m let $\mathscr{K}(m)$ be the number $\min\{i \in \mathbb{N} \mid \varphi_i(0)\downarrow = m\}$. ⊛

The function $\mathscr{K}$ is total and for any $e \in \mathbb{N}$ there are finitely many m's which satisfy $\mathscr{K}(m) \leqslant e$. The following is Lemma 7 of [1].

Lemma 3.2 *(Uncomputability of Complexity). There is no computable function f which satisfies the inequality $\mathscr{K}\big(f(m)\big) > m$ for all $m \in \mathbb{N}$.*

Proof. If there were such a computable function f, then by Kleene's second recursion theorem there would exist some e such that $\varphi_e(x) = f(e)$ and so, in particular, $\varphi_e(0) = f(e)$ which implies $\mathscr{K}\big(f(e)\big) \leqslant e$; a contradiction. ⊠

So, $\mathscr{K}$ is not computable, since otherwise $f(x) = \min\{y \mid \mathscr{K}(y) > x\}$, which satisfies $\forall x : \mathscr{K}\big(f(x)\big) > x$, would be computable.

Theorem 3.3 *(Chaitin's Theorem). For any consistent RE theory T which contains Q there exists a constant $c_T \in \mathbb{N}$ such that for any $e \geqslant c_T$ and any $w \in \mathbb{N}$ we have $T \nvdash$ "$\mathscr{K}(w) > e$".*

Proof. If not, then for any given $m \in \mathbb{N}$ there exists some $e \geqslant m$ and some w such that $T \vdash$ "$\mathscr{K}(w) > e$". Let us note that if $T \vdash$ "$\mathscr{K}(w) > e$" for a consistent $T \supseteq Q$ then $\mathscr{K}(w) > e$, since otherwise, if $\mathscr{K}(w) \leqslant e$, the true Σ_1-sentence "$\mathscr{K}(w) \leqslant e$" would be provable in Q (and so in T) which contradicts the consistency of T. Now, for a given m we can, by an algorithmic proof search in T, find some $e \geqslant m$ and w such that $T \vdash$ "$\mathscr{K}(w) > e$" (and so $\mathscr{K}(w) > e$); our assumption guarantees the termination of this algorithm for any input m. Let $f(m)$ be one of those w's; then we have $\mathscr{K}\big(f(m)\big) > e \geqslant m$ which contradicts Lemma 3.2. ⊠

This is an incompleteness theorem since for any c there are cofinitely many w's with $\mathscr{K}(w) > c$. So, for a given T which is consistent and RE and contains Q there are cofinitely many w's such that the true sentences "$\mathscr{K}(w) > c_T$" are unprovable in T. As for the constructivity of this proof, the good news is that a constant c_T which satisfies Chaitin's Theorem 3.3 can be algorithmically constructed from T.

Theorem 3.4 *(Computing a Chaitin Constant). For a given consistent and* RE *extension T of Q one can algorithmically construct a constant c_T such that for all $e \geqslant c_T$ and all w, we have $T \nvdash$ "$\mathscr{K}(w) > e$".*

Proof. Given a description of a consistent, Σ_1-complete and RE theory T the following can be done algorithmically. Define $\hbar(x)$ to be the first ordered pair $\langle a, b\rangle$ such that the proof search algorithm of T shows up (a proof of) the sentence "$\mathscr{K}(a) > b \geqslant x$" (so, $T \vdash$ "$\mathscr{K}(a) > b \geqslant x$"). This is (a partially) computable (function) and an index of it can be calculated from (a description of) T. By Kleene's second recursion theorem there exists a constant c such that $\varphi_c(y) = \hbar_1(c)$, where $\hbar_1(x)$ is the first component of the ordered pair $\hbar(x)$. The constant c can be computed from T (since Kleene's second recursion theorem is itself constructive); let us denote it by c_T. Now, we show that for no $b \geqslant c_T$ and no a can $T \vdash$ "$\mathscr{K}(a) > b$" hold. If there exists such a and b then we have $T \vdash$ "$\mathscr{K}(a) > b \geqslant c_T$". If $\langle a, b\rangle$ is the first ordered pair such that "$\mathscr{K}(a) > b \geqslant c_T$" appears in the above mentioned proof search algorithm of T, then $\hbar(c_T) = \langle a, b\rangle$ and so $\varphi_{c_T}(0) = \hbar_1(c_T) = a$. Thus, $\mathscr{K}(a) \leqslant c_T$, and by the Σ_1-completeness of T we have $T \vdash$ "$\mathscr{K}(a) \leqslant c_T$". But from $T \vdash$ "$\mathscr{K}(a) > b \geqslant c_T$" we have $T \vdash$ "$\mathscr{K}(a) > c_T$", contradicting the consistency of T. ⊠

Unfortunately, by Lemma 3.2 one cannot calculate a w with $\mathscr{K}(w) > c_T$ given c_T for a theory T. Otherwise one could get a constructive version of Chaitin's proof: Given a consistent and RE theory $T \supseteq Q$ one calculates c_T and finds some w with $\mathscr{K}(w) > c_T$; then "$\mathscr{K}(w) > c_T$" is a true sentence which is not provable in T. It is actually known that Chaitin's proof is not constructive; see e.g. [16, page 1394] or [23, page 95].

Theorem 3.5 *(Non-Constructivity of Chaitin's Proof). There is no algorithm such that for a given consistent and* RE *extension T of Q can compute some w_T such that $\mathscr{K}(w_T) > c_T$ holds, where c_T is a Chaitin constant as in Theorem 3.4.*

Proof. If such a w_T were computable from T, then the theory $T_\infty = \bigcup_{i\in\mathbb{N}} T_i$ would be RE where $T_0 = Q$ and inductively $T_{i+1} = T_i +$ "$\mathscr{K}(w_{T_i}) > c_{T_i}$" are defined by iterating the computation procedure. The theory T_∞ is also consistent (indeed, sound) and contains Q, so by Chaitin's Theorem 3.3 there should exist some constant c_{T_∞} such that for no w can we have the deduction $T_\infty \vdash$ "$\mathscr{K}(w) > c_{T_\infty}$". But this is a contradiction because we have $c_{T_i} < c_{T_{i+1}}$ and also $c_{T_i} < c_{T_\infty}$ for all $i \in \mathbb{N}$. ⊠

An Alternative Proof. Albert Visser suggested the following argument as another proof of Theorem 3.5: Since the sequence $\{c_{T_i}\}_{i\in\mathbb{N}}$ is strictly increasing, we have $c_{T_m} \geqslant m$ for any $m \in \mathbb{N}$. Now, $\forall m \in \mathbb{N} : \mathscr{K}(w_{T_m}) > c_{T_m} \geqslant m$ would contradict Lemma 3.2 if w_T were computable from T. ⊠

Remark 3.6. Albert Visser noted that Theorems 3.4 and 3.5 amusingly imply Lemma 3.2, since if there were a computable (total) function f with $\forall m \in \mathbb{N} : \mathscr{K}\big(f(m)\big) > m$ then one could take w_T as $f(c_T)$. ⊛

The true unprovable sentences "$\mathscr{K}(w) > e$" (for $e \geqslant c_T$) are also independent when T is a (Σ_1-)sound theory: If $T \vdash$ "$\mathscr{K}(w) \leqslant e$" then the Σ_1-sentence $\mathscr{K}(w) \leqslant e$ has to be true, a contradiction. So, we restate Chaitin's Theorem as

Corollary 3.7 *(Chaitin's Theorem, restated). Let T be a Σ_1-sound and* RE *theory such that $T \supseteq \mathsf{Q}$. There exists some c_T (which is computable from T) such that for any $e \geqslant c_T$ there are cofinitely many w's such that "$\mathscr{K}(w) > e$" is independent from T.* $\boxtimes$

For having a Rosserian version of Chaitin's Theorem we will replace the assumption of the "Σ_1-soundness" (of T) in Corollary 3.7 with (its simple) "consistency". For doing that we need the following version of the Pigeonhole Principle in Q (which holds in R as well).

Lemma 3.8 *(A Pigeonhole Principle). For any $k \in \mathbb{N}$ we have*

$$\mathsf{Q} \vdash \forall z_0, \cdots, z_k \Big(\bigwedge_{0 \leqslant i \leqslant k} z_i < \overline{k} \quad \longrightarrow \quad \bigvee_{0 \leqslant i,j \leqslant k}^{i \neq j} z_i = z_j \Big).$$

Proof. This can be proved by induction (in the metalanguage) on k: for $k = 0$ it suffices to note that $\mathsf{Q} \vdash \forall z \neg(z < 0)$ and for the induction step it suffices to use the derivation $\mathsf{Q} \vdash \forall z (z < \overline{k+1} \rightarrow z < \overline{k} \vee z = \overline{k})$; cf. [21, page 73]. $\boxtimes$

Theorem 3.9 *(Rosserian form of Chaitin's Theorem). For any consistent* RE *extension T of Q there is a constant c_T (which is computable from T) such that for any $e \geqslant c_T$ there are cofinitely many w's such that "$\mathscr{K}(w) > e$" is independent from T.*

Proof. By Chaitin's Theorem 3.3 there exists a constant c_T (which is computable from T) such that for any $e \geqslant c_T$ there are cofinitely many w's such that "$\mathscr{K}(w) > e$" is true but unprovable in T. Fix an $\mathsf{e} \geqslant c_T$. For no w can $T \vdash$ "$\mathscr{K}(w) > \mathsf{e}$" hold, and $T \vdash$ "$\mathscr{K}(w) \leqslant \mathsf{e}$" can hold for at most $(\mathsf{e}+1)$-many w's: if for some distinct $\mathsf{w}_0, \mathsf{w}_1, \ldots, \mathsf{w}_{\mathsf{e}+1}$, the derivations $T \vdash$ "$\mathscr{K}(\mathsf{w}_i) \leqslant \mathsf{e}$" hold ($i = 0, 1, \ldots, \mathsf{e}+1$) then $T \vdash \exists z_0, z_1, \ldots, z_{\mathsf{e}+1} \big(\bigwedge_{i=0}^{\mathsf{e}+1} [z_i \leqslant \overline{\mathsf{e}} \wedge \varphi_{z_i}(0){\downarrow} = \mathsf{w}_i] \big)$ and so

$$T \vdash \exists z_0, z_1, \ldots, z_{\mathsf{e}+1} \big(\bigwedge\nolimits_{0 \leqslant i \leqslant \mathsf{e}+1} [z_i < \overline{\mathsf{e}} + 1] \wedge \bigwedge\nolimits_{0 \leqslant i,j \leqslant \mathsf{e}+1}^{i \neq j} [z_i \neq z_j] \big)$$

which contradicts Lemma 3.8 (for $k = \mathsf{e}+1$). Thus, for cofinitely many w's we have both $T \nvdash$ "$\mathscr{K}(w) > \mathsf{e}$" and $T \nvdash$ "$\mathscr{K}(w) \leqslant \mathsf{e}$". $\boxtimes$

Martin Davis [5] calls Chaitin's Theorem "a dramatic extension of Gödel's incompleteness theorem". We saw that this theorem as presented in Corollary 3.7 can be hardly considered an extension of Gödel's incompleteness theorem, as Gödel's proof is constructive while Chaitin's is not (Theorem 3.5). The Rosserian form of Chaitin's Theorem as presented in Theorem 3.9 could be considered as an extension of Gödel's and Chaitin's theorems in a sense, even though, it is not any more extension than Rosser's own [17]; let us also note that Rosser's proof is constructive (while the proof of Theorem 3.9 is not).

4. The proof of Boolos for Gödel's incompleteness theorem

Jon Barwise calls it "a very lovely proof of Gödel's Incompleteness Theorem, probably the deepest single result about the relationship between computers and mathematics", and mentions that it is "the most straightforward proof of this result that I have ever seen".[1] After its first appearance in [2] this proof was discussed, extended and studied in e.g. [9,10,18,15,20,11].

[1] J. Barwise, "Editorial Notes: This Month's Column", *Notices of the American Mathematical Society*, vol 36, no. 4 (1989), page 388.

Notation 4.1 *(Arithmetization).* For an RE theory T denote the provability predicate of T by $\mathit{Pr}_T(x)$; so $\mathit{Con}(T) = \neg\mathit{Pr}_T(\bot)$ is the consistency statement of T. Suppose that the variables are $x, x', x'', x''', \cdots$ whose lengths are $1, 2, 3, 4, \cdots$, respectively. ⊛

So, for any $k \in \mathbb{N}$ there are at most finitely many formulas with length k.

Definition 4.2 *(Formalizing Berry's Paradox).* For a formula $\psi(x_1, \cdots, x_m)$ with the shown (possibly empty) set of free variables ($m \geqslant 0$) and number n, let $\mathsf{D}(\psi, n)$ be the (Gödel code of) $\forall x[\psi(x, \cdots, x) \leftrightarrow x = \overline{n}]$. The number n is definable in T by the formula ψ when $\mathit{Pr}_T\big(\mathsf{D}(\psi, n)\big)$ holds.

Let $\mathit{Def}_T^{<z}(y) = \exists x\big[\mathit{len}(x) < z \wedge \mathit{Pr}_T\big(\mathsf{D}(x, y)\big)\big]$, where $\mathit{len}(x)$ denotes the length of (the formula with Gödel code) x. The formula $\mathit{Def}_T^{<z}(y)$ states that "there exists a formula $\psi(x_1, \cdots, x_m)$ whose length is smaller than z such that the deduction $T \vdash \forall x[\psi(x, \cdots, x) \leftrightarrow x = \overline{y}]$ holds", or informally "the number y is definable in T by a formula with length less than z".

Let $\mathit{Berry}_T^{<v}(u) = \neg\mathit{Def}_T^{<v}(u) \wedge \forall y < u\ \mathit{Def}_T^{<v}(y)$, meaning that "$u$ is the least number not definable by a formula with length less than v".

Let ℓ_T be the length of the formula $\mathit{Berry}_T^{<x'}(x)$ and let $\mathit{Boolos}_T(x)$ be the formula

$$\exists x'\big[x' = \overline{5} \cdot \overline{\ell_T} \wedge \mathit{Berry}_T^{<x'}(x)\big].$$

Let b_T be the least number not definable by a formula with length less than $5\ell_T$. ⊛

Theorem 4.3 *(Boolos' Theorem). For any consistent and* RE *extension T of Q, the sentence $\mathit{Boolos}_T(\overline{b_T})$ is (true but) unprovable in T.*

Proof. First we show that $\mathsf{Q} \vdash \forall u, v[\mathit{Berry}_T^{<v}(\overline{n}) \wedge \mathit{Berry}_T^{<v}(u) \rightarrow \overline{n} = u]$ holds for any $n \in \mathbb{N}$. Reason inside Q: if for some u, v we have (a) $\mathit{Berry}_T^{<v}(\overline{n})$ and (b) $\mathit{Berry}_T^{<v}(u)$ then (a') $\neg\mathit{Def}_T^{<v}(\overline{n})$, (a”) $\forall y < \overline{n}\mathit{Def}_T^{<v}(y)$, (b') $\neg\mathit{Def}_T^{<v}(u)$ and (b”) $\forall y < u\mathit{Def}_T^{<v}(y)$ hold. Now, by $u \leqslant \overline{n} \vee \overline{n} \leqslant u$, if $u \neq \overline{n}$ then either $u < \overline{n}$ or $\overline{n} < u$ holds. In the former case we have a contradiction between (a”) and (b'), and in the latter case we have a contradiction between (a') and (b”). Therefore, $\overline{n} = u$. Now, assume (for the sake of contradiction) that $T \vdash \mathit{Boolos}_T(\overline{b_T})$. Then $T \vdash \forall u, v[\mathit{Berry}_T^{<v}(\overline{n}) \wedge \mathit{Berry}_T^{<v}(u) \rightarrow \overline{n} = u]$, shown above, implies the deduction $T \vdash \forall x[\mathit{Boolos}_T(x) \leftrightarrow x = \overline{b_T}]$. Thus, b_T is definable in T by the formula $\mathit{Boolos}_T(x)$ whose length is less than $\ell_T + \mathit{len}(\overline{5} \cdot \overline{\ell_T}) + 9 = 4\ell_T + 26 < 5\ell_T$ (since, for any m, the term $\overline{m} = \mathsf{s}(\cdots(\mathsf{s}(0))\dots)$ [m-times s] has length $3m + 1$). So, the Σ_1-sentence $\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b_T})$ is true, thus provable in Q; whence $T \vdash \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b_T})$. On the other hand $T \vdash \mathit{Boolos}_T(\overline{b_T})$ implies that $T \vdash \neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b_T})$, contradicting the consistency of T. ⊠

The formula $\mathit{Boolos}_T(\overline{b_T})$ is not Π_1; however, the following modification from [9] proves a Π_1-incompleteness.

Theorem 4.4 *(Boolos' Theorem, modified). For any consistent and* RE *extension T of Q, the true Π_1-sentence $\neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b_T})$ is unprovable in T.*

Proof. Assume, to the contrary, that $T \vdash \neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b_T})$. Since any number y less than b_T is definable by a formula with length less than $5\ell_T$ then the Σ_1-sentence $\forall y < b_T\ \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(y)$ is true and thus provable in T. Therefore, $\neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b_T}) \wedge \forall y < b_T\ \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(y)$ is provable in T and so $T \vdash \mathit{Boolos}_T(\overline{b_T})$, contradicting Theorem 4.3. ⊠

Even though ℓ_T is computable from T, below we show that one cannot calculate b_T.

Theorem 4.5 *(Non-Constructivity of Boolos' Proof). There is no algorithm such that for a given consistent and* RE *extension T of Q can compute b_T.*

Proof. Assume that b_T is computable from T, and let $T_0 = \mathsf{Q}$ and inductively $T_{j+1} = T_j + \neg\mathsf{Def}_{T_j}^{<\overline{5}\cdot\overline{\ell_{T_j}}}(\overline{\mathsf{b}_{T_j}})$. Define the function $\hbar(n)$, for any $n \in \mathbb{N}$, to be the greatest m with $\forall j < m : \mathsf{len}(\text{“}\boldsymbol{\varphi}_j(0){\downarrow} = x\text{”}) < n$. This is a computable and non-decreasing function; also $\lim_n \hbar(n) = \infty$. So, from $\lim_j \ell_{T_j} = \infty$ we have $\lim_j \hbar(5\ell_{T_j}) = \infty$. Therefore, for any (given) x one can compute some $\iota(x)$ such that $\hbar(5\ell_{T_{\iota(x)}}) > x$. The proof will be complete when we show that $\mathscr{K}(\mathsf{b}_{T_j}) \geqslant \hbar(5\ell_{T_j})$ holds for any j: Because, by the computability of b_{T_j} from j, we will have a computable function $x \mapsto \mathsf{b}_{T_{\iota(x)}}$ which satisfies $\forall x : \mathscr{K}\big(\mathsf{b}_{T_{\iota(x)}}\big) \geqslant \hbar(5\ell_{T_{\iota(x)}}) > x$ contradicting Lemma 3.2. For showing that $\mathscr{K}(\mathsf{b}_{T_j}) \geqslant \hbar(5\ell_{T_j})$ holds for any j, we show more generally that for any u, v if $\neg\mathsf{Def}_T^{<v}(u)$ holds, for some consistent $T \supseteq \mathsf{Q}$, then $\mathscr{K}(u) \geqslant \hbar(v)$: If, to the contrary, we have $\mathscr{K}(u) < \hbar(v)$ then there exists some j such that (1) $j < \hbar(v)$ and (2) $\boldsymbol{\varphi}_j(0){\downarrow} = u$. By (2) the number u is definable by the formula “$\boldsymbol{\varphi}_j(0){\downarrow} = x$” in Q (and so in T), and by (1) the length of the formula “$\boldsymbol{\varphi}_j(0){\downarrow} = x$” is less than v; so $\mathsf{Def}_T^{<v}(u)$ should hold, a contradiction. ⊠

Of course, when T is Σ_1-sound then $\neg\mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{\mathsf{b}_T})$ is independent from T. Also $\mathsf{Boolos}_T(\overline{\mathsf{b}_T})$ is independent from T: Because if $T \vdash \neg\mathsf{Boolos}_T(\overline{\mathsf{b}_T})$ then $T \vdash \neg\mathsf{Berry}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{\mathsf{b}_T})$ and so $T \vdash \forall y < \overline{\mathsf{b}_T}\ \mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(y) \rightarrow \mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{\mathsf{b}_T})$. But $\forall y < \overline{\mathsf{b}_T}\ \mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(y)$, being a true Σ_1-sentence, is provable in T. Whence, we have $T \vdash \mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{\mathsf{b}_T})$, a contradiction. However, we show in the following theorem that if T is not Σ_1-sound then $\mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{\mathsf{b}_T})$, and so $\neg\mathsf{Boolos}_T(\overline{\mathsf{b}_T})$, could be provable in T. For the following theorem to make sense we note that for any theory U satisfying the following conditions

(i) $U \nvdash \mathsf{Con}(U)$, i.e., Gödel's Second Incompleteness Theorem holds for U;
(ii) $U \vdash \mathsf{Con}(U + \psi) \rightarrow \mathsf{Con}(U)$, for any ψ;

there exists a consistent theory $S \supseteq U$ such that $S + \mathsf{Con}(S)$ is not consistent: The theory $S = U + \neg\mathsf{Con}(U)$ is consistent by (i), and $S \vdash \neg\mathsf{Con}(S)$ because $S \vdash \neg\mathsf{Con}(U)$ by the definition of S and $S \vdash \neg\mathsf{Con}(U) \rightarrow \neg\mathsf{Con}(S)$ by (ii).

One example for a theory that satisfies the conditions (i) and (ii) above, and also (iii) in Theorem 4.6 and (iv) in Theorem 4.7 below, is Peano's Arithmetic. This arithmetic is indeed too strong and the finitely axiomatizable theory $\mathrm{I}\Sigma_1$ (see [7]) satisfies the conditions (i), (ii), (iii) and (iv). Even the weaker theories $\mathrm{I}\Delta_0 + \Omega_1$ (see [25]) and S_2^1 (see [3]) are strong enough to satisfy them.

Theorem 4.6 *(Boolos' Proof is not Rosserian). Suppose that a consistent and* RE *extension T of Q satisfies the following condition for any formula ψ:*

(iii) $T \vdash \mathsf{Pr}_T(\bot) \rightarrow \mathsf{Pr}_T(\psi)$.

If $T + \mathsf{Con}(T)$ is $\underline{in}$consistent then for any $b \in \mathbb{N}$ we have $T \vdash \mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$, and so $T \vdash \neg\mathsf{Boolos}_T(\overline{b})$.

Proof. If $T \vdash \neg\mathsf{Con}(T)$ then $T \vdash \mathsf{Pr}_T(\bot)$ and so $T \vdash \mathsf{Pr}_T(\psi)$, for any ψ, by the condition (iii). In particular, if ψ is a formula with length less than $5\ell_T$ (for example $\mathsf{Berry}^{<x'}(x)$) then $T \vdash \mathsf{Pr}_T\big(\mathsf{D}(\psi, b)\big)$ and so for any arbitrary number b we have $T \vdash \exists x\big[\mathsf{len}(x) < \overline{5} \cdot \overline{\ell_T} \wedge \mathsf{Pr}_T\big(\mathsf{D}(x, b)\big)\big]$. Thus $T \vdash \mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$, whence $T \vdash \neg\mathsf{Berry}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$ and $T \vdash \neg\mathsf{Boolos}_T(\overline{b})$. ⊠

So, if $T + \mathsf{Con}(T)$ is not consistent, then $\neg\mathsf{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$ is not independent from the theory T (neither is $\mathsf{Boolos}_T(\overline{b})$) for any b. However, if $T + \mathsf{Con}(T)$ is consistent, then a variant of Boolos' proof can go through (cf. [10, Theorem 7.2]) as is shown in the following theorem.

Theorem 4.7 *(Boolos' Theorem, restated). If an* RE *extension T of Q satisfies the following condition for any $m, n, k \in \mathbb{N}$,*

(iv) $T \vdash \mathit{Pr}_T\big(D(k,m)\big) \wedge \mathit{Pr}_T\big(D(k,n)\big) \wedge \overline{m} \neq \overline{n} \to \neg \mathit{Con}(T)$,

and the theory $T + \mathit{Con}(T)$ is consistent, then there exists some $b \in \mathbb{N}$ such that $\neg \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$, and also $\mathit{Boolos}_T(\overline{b})$, is independent from T.

Proof. First we show that there exists some a such that $T \nvdash \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{a})$. If not, then for any i we have $T \vdash \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{i})$. Let $\Bbbk$ be a fixed number greater than the maximum Gödel codes of formulas ϕ with $\mathit{len}(\phi) < 5\ell_T$. So, for any $i \leqslant \Bbbk$ we have $T \vdash \exists z < \Bbbk \; \mathit{Pr}_T\big(\mathsf{D}(z,i)\big)$. By Lemma 3.8 there exists some $i < j \leqslant \Bbbk$ and some $\ell < \Bbbk$ such that $T \vdash \mathit{Pr}_T\big(\mathsf{D}(\ell,i)\big) \wedge \mathit{Pr}_T\big(\mathsf{D}(\ell,j)\big)$. Now (iv) implies that $T \vdash \neg\mathit{Con}(T)$, a contradiction. Let b be the minimum of those a's with $T \nvdash \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{a})$. So, $T \vdash \forall z < \overline{b} \, \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(z)$. Now we show that $T \nvdash \neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$: If not ($T \vdash \neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$) then $T \vdash \mathit{Berry}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$ or equivalently, $T \vdash \mathit{Boolos}_T(\overline{b})$. So, b is definable in T by a formula with length less than $5\ell_T$ (see the proof of Theorem 4.3) whence $\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$ is true thus provable in T; a contradiction. Therefore, we showed that $T \nvdash \mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$ and $T \nvdash \neg\mathit{Def}_T^{<\overline{5}\cdot\overline{\ell_T}}(\overline{b})$ (also $T \nvdash \mathit{Boolos}_T(\overline{b})$ and $T \nvdash \neg\mathit{Boolos}_T(\overline{b})$). ⊠

Thus, the consistency of $T + \mathit{Con}(T)$ is an optimal (indeed, necessary and sufficient) condition for the independence of a Boolos sentence from T.

5. Concluding remarks

The following table summarizes some of the new and old results in this paper:

Proof	Constructive	Rosser Property
GÖDEL (1931) [6]	✓	✗ [8]
ROSSER (1936) [17]	✓	✓
KLEENE$_1$ (1936) [12]	✓	✗ Theorem 2.3
KLEENE$_2$ (1950) [13]	✓	✓
CHAITIN (1971) [4]	✗ [16,23], Theorem 3.5	✓ Theorem 3.9
BOOLOS (1989) [2]	✗ Theorem 4.5	✗ Theorem 4.6

Let us note that for the constructivity of a proof, usually, no new argument is needed as a computational procedure could often be seen from the proof. But the non-constructivity of a proof (as in the case of Chaitin's and Boolos' proofs) should be proved; proving the non-constructivity (the non-existence of any algorithm) is usually harder than showing the constructivity (the existence of an algorithm). So is having the Rosser property of a proof. Other than Rosser's proof and Kleene's symmetric theorem (1950) Chaitin's proof has the Rosser property. The non-Rosserian proofs of Gödel and Boolos need the consistency of $T + \mathit{Con}(T)$ for the independence of their true but unprovable sentences, and this condition, $\mathit{Con}\big(T + \mathit{Con}(T)\big)$, is optimal (for the independence of that sentences).

Acknowledgements

The authors warmly thank Professor Albert Visser for carefully reading the whole paper and for his comments which greatly improved the results of the paper and its presentation and notation. As noted inside the text, the alternative proof of Theorem 3.5 and Remark 3.6 are his ideas, and the proof of Theorem 4.3 has been corrected by his suggestions.

References

[1] L.D. Beklemishev, Gödel incompleteness theorems and the limits of their applicability, I, Russian Math. Surveys 65 (2010) 857–899, https://doi.org/10.1070/RM2010v065n05ABEH004703.

[2] G. Boolos, A new proof of the Gödel incompleteness theorem, Notices Amer. Math. Soc., vol. 36, ISBN 9780674537675, 1989, pp. 388–390;
Reprinted in G. Boolos, Logic, Logic, and Logic, Harvard University Press, 1998, pp. 383–388.

[3] S. Buss, Bounded Arithmetic, Bibliopolis, Naples, Italy, 1986, https://www.math.ucsd.edu/~sbuss/ResearchWeb/BAthesis/.

[4] G.J. Chaitin, Computational complexity and Gödel's incompleteness theorem, SIGACT News 9 (1971) 11–12, https://doi.org/10.1145/1247066.1247068.

[5] M. Davis, "What is a computation?", in: L.A. Steen (Ed.), Mathematics Today, Twelve Informal Essays, Springer, 1978, pp. 241–267.

[6] K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I, Monatsh. Math. Phys. 38 (1931) 173–198, https://doi.org/10.1007/BF01700692;
Translated as "On Formally Undecidable Propositions of Principia Mathematica and Related Systems, I", in: S. Feferman, et al. (Eds.), Kurt Gödel Collected Works, Volume I: Publications 1929–1936, Oxford University Press, 1986, pp. 135–152.

[7] P. Hájek, P. Pudlák, Metamathematics of First-Order Arithmetic, 2nd. print, Springer-Verlag, ISBN 9783540636489, 1998, http://projecteuclid.org/euclid.pl/1235421926.

[8] D. Isaacson, "Necessary and sufficient conditions for undecidability of the Gödel sentence and its truth", in: D. DeVidi, et al. (Eds.), Logic, Mathematics, Philosophy: Vintage Enthusiasms, Springer, ISBN 9789400702134, 2011, pp. 135–152.

[9] M. Kikuchi, A note on Boolos' proof of the incompleteness theorem, MLQ Math. Log. Q. 40 (1994) 528–532, https://doi.org/10.1002/malq.19940400409.

[10] M. Kikuchi, K. Tanaka, On formalization of model-theoretic proofs of Gödel's theorems, Notre Dame J. Form. Log. 35 (1994) 403–412, https://doi.org/10.1305/ndjfl/1040511346.

[11] M. Kikuchi, T. Kurahashi, H. Sakai, On proofs of the incompleteness theorems based on Berry's paradox by Vopěnka, Chaitin, and Boolos, MLQ Math. Log. Q. 58 (2012) 307–316, https://doi.org/10.1002/malq.201110067.

[12] S.C. Kleene, General recursive functions of natural numbers, Math. Ann. 112 (1936) 727–742, https://doi.org/10.1007/BF01565439.

[13] S.C. Kleene, A symmetric form of Gödel's theorem, Indag. Math. (N.S.) 12 (1950) 244–246 (issn:0019–3577).

[14] S.C. Kleene, Introduction to Metamathematics, North-Holland, ISBN 9780720421033, 1952.

[15] H. Kotlarski, The incompleteness theorems after 70 years, Ann. Pure Appl. Logic 126 (2004) 125–138, https://doi.org/10.1016/j.apal.2003.10.012.

[16] M. van Lambalgen, Algorithmic information theory, J. Symbolic Logic 54 (1989) 1389–1400, https://doi.org/10.1017/S0022481200041153.

[17] B. Rosser, Extensions of some theorems of Gödel and Church, J. Symbolic Logic 1 (1936) 87–91, https://doi.org/10.2307/2269028.

[18] D.K. Roy, The shortest definition of a number in Peano arithmetic, MLQ Math. Log. Q. 49 (2003) 83–86, https://doi.org/10.1002/malq.200310006.

[19] S. Salehi, Gödel's incompleteness phenomenon—computationally, Philos. Sci. 18 (2014) 23–37, https://doi.org/10.4000/philosophiascientiae.968.

[20] G. Serény, Boolos-style proofs of limitative theorems, MLQ Math. Log. Q. 50 (2004) 211–216, https://doi.org/10.1002/malq.200310091.

[21] P. Smith, An Introduction to Gödel's Theorems, 2nd edition, Cambridge University Press, ISBN 9781107606753, 2013.

[22] C. Smoryński, "The incompleteness theorems", in: J. Barwise (Ed.), Handbook of Mathematical Logic, North-Holland, ISBN 9780444863881, 1977, pp. 821–865.

[23] J.C. Stillwell, Roads to Infinity: The Mathematics of Truth and Proof, A K Peters/CRC Press, ISBN 9781568814667, 2010.

[24] A. Tarski, A. Mostowski, R.M. Robinson, Undecidable Theories, North-Holland, ISBN 9780486477039, 1953, reprinted by Dover Publications, 2010.

[25] A.J. Wilkie, J.B. Paris, On the scheme of induction for bounded arithmetic formulas, Ann. Pure Appl. Logic 35 (1987) 261–302, https://doi.org/10.1016/0168-0072(87)90066-2.