Notre Dame Journal of Formal Logic
Volume XIV, Number 3, July 1973
NDJFAM

INCOMPLETENESS THEOREM VIA WEAK DEFINABILITY OF TRUTH: A SHORT PROOF

GIORGIO GERMANO

Introduction According to Tarski [9] no consistent theory T in which all recursive functions are definable allows definability of the truth of its own sentences, in the sense that there exists no formula Θ (with at most one free variable) such that for all sentences Φ

$$(\Theta(\mathsf{D}_{g(\Phi)}) \longleftrightarrow \Phi) \in T,$$

where $g(\Phi)$ is the Gödel-number of Φ and $\mathsf{D}_{g(\Phi)}$ is the digit representing it. We shall refer to the definability of truth in this sense as "strong definability of truth." Myhill [8] defines a system S which allows definability of its own truth in the sense that there is a Θ such that for all Φ

$$\Theta(\mathsf{D}_{g(\Phi)}) \in S \textit{ iff } \Phi \in S.$$

We shall refer to the definability of truth in this sense as "weak definability of truth," because if truth is definable in the first sense, it follows that truth is definable in the later one.

In [2] Germano, solving a problem which in [1] Germano has left open, proves that the weak truth definability is a property of every recursively enumerable arithmetic in which all recursive functions are definable.

In [3] Germano gives a strong formulation of the incompleteness theorem (concerning every recursively enumerable arithmetic in which the elementary functions, i.e., the functions of the class $\mathcal{E}^2$ of Grzegorczyk [5], are definable) by comparing opportune formulations of the theorem on weak definability of truth and of the theorem on strong definability of truth. The present note gives a proof of the incompleteness theorem in the same strong formulation as in Germano [3], using only the theorem on weak definability of truth as announced in Germano [4]. The proof obtained in this way is the shortest direct proof of the incompleteness theorem known to the author and it is characterized by the fact that it mirrors step by step the construction of the *liar*'s paradox, which will be discussed later. A future paper will treat the possibility of extending the incompleteness

Received May 10, 1972

theorem to every arithmetic in which only a few of the functions of the class $\mathcal{E}^2$ of Grzegorczyk [5] are definable. A problem, which should be interesting with regards to applications, is the following: can the incompleteness theorem be extended to every arithmetic in which only functions of the class $\mathcal{E}^0$ are definable?

Nomenclature In order to obtain a more compact exposition, we consider a language (of order $n \geqslant 1$) constructed using three (denumerable) sets of individual symbols: free individual variables $a_0, a_1, \ldots$, bound individual variables $x_0, x_1, \ldots$ and individual constants (as digits) $\mathsf{D}_0, \mathsf{D}_1, \ldots$. Furthermore, the equality symbol $=$, the negation $\neg$, the existential quantifier of the first order and at most denumerably many other symbols shall be used. The considerations below apply nevertheless to many other languages, e.g., to the languages of Tarski [10].

Let f be any injective mapping of the set of the introduced symbols into the set of positive integers, such that $f(\mathsf{D}_n) = hn + 1$, where h is a constant. For any finite sequence of symbols $s_0 \ldots s_n$ let

$$\mathsf{g}(s_0 \ldots s_n) = \prod_{i \in n} p_i^{f(s_i)} ,$$

where p_i is the i'th prime number. Now consider the function

$$d(n) = \prod_{i \leqslant \mathsf{l}(n)} p_i^{a(i)} ,$$

where $\mathsf{l}(n)$ is the length of n and $a(i) = \begin{cases} hn + 1 \text{ for } \mathsf{exp}_i(n) = f(a_0) \\ \mathsf{exp}_i(n) \text{ otherwise.} \end{cases}$

The function d is obviously elementary. Let E be any finite sequence of symbols and $E(\mathsf{D}_n)$ the sequence obtained by replacing every occurrence of the free individual variable a_0 in E by D_n. Then obviously

$$d(\mathsf{g}(E)) = \mathsf{g}(E(\mathsf{D}_{\mathsf{g}(E)})),$$

i.e., the function d is a diagonal function.

A theory shall be a set of sentences (well formed formulas without free variables) closed with respect to the inference rules of the predicate calculus of the first order with identity. A theory will be said to be an arithmetic iff it is satisfied by a realization whose universe is the set of negative integers and which interprets every D_n as n. A theory will be said to be recursively enumerable iff its image under g is recursively enumerable.

For the sake of compactness we will use in the metalanguage the connectives "$\Longrightarrow$" and "$\Longleftrightarrow$", the quantifiers "$\forall$" and "$\exists$" and the following letters: "T" representing theories, "Φ" representing sentences, "Θ" representing well formed formulas with no other free variable than a_0 and "θ", "δ" representing terms (individual-to-individual functional expressions) with no other free variable than a_0.

We will say that truth can be weakly defined in T iff

$$\exists \Theta \forall \Phi \, (\Theta(\mathsf{D}_{\mathsf{g}(\Phi)}) \in T \Longleftrightarrow \Phi \in T)$$

We will say that the function f can be defined in T iff

$$\exists \theta \forall n \ \theta(\mathsf{D}_n) = \mathsf{D}_{f(n)} \in T.$$

Also this less general form of definability of functions (see Mostowski [7]) is used in order to obtain a more compact exposition. Nevertheless we could apply the same method also using a more general form as, e.g., in Tarski [10], p. 46, Theorem 1.

Theorem *If* (a) *T is an arithmetic,* (b) *T is recursively enumerable, and* (c) *every elementary function is definable in T, then* (A) *truth can be weakly defined in T, and* (B) *T is incomplete.*

Proof: By (b) there is an elementary function f such that

(1) $\Phi \in T \Longleftrightarrow \exists n\, f(n) = \mathrm{g}(\Phi)$.

By (c) there is a θ such that

(2) $\forall n\, \theta(\mathsf{D}_n) = \mathsf{D}_{f(n)} \in T$.

By (2)

(3) $\exists n\, \mathrm{g}(\Phi) = f(n) \Longrightarrow \forall x_0 \ \theta(x_0) = \mathsf{D}_{\mathrm{g}(\Phi)} \in T$

and, by (a) and (2)

(4) $\forall x_0 \ \theta(x_0) = \mathsf{D}_{\mathrm{g}(\Phi)} \in T \Longrightarrow \exists n\, f(n) = \mathrm{g}(\Phi)$.

From (1), (3) and (4) we get

$$\text{(A)} \quad \forall x_0 \ \theta(x_0) = \mathsf{D}_{\mathrm{g}(\Phi)} \in T \Longleftrightarrow \Phi \in T.$$

As d is elementary, by (c) there is a δ such that

(5) $\forall n \delta(\mathsf{D}_n) = \mathsf{D}_{d(n)} \in T$.

Therefore, for $n := \mathrm{g}(\neg \forall x_0\, \theta(x_0) = \delta)$ and $e := d(n) = \mathrm{g}(\neg \forall x_0\, \theta(x_0) = \delta(\mathsf{D}_n))$

(6) $\forall x_0\, \theta(x_0) = \mathsf{D}_e \in T \Longleftrightarrow \neg \forall x_0\, \theta(x_0) = \delta(\mathsf{D}_n) \in T$, by (A)
$\Longleftrightarrow \neg \forall x_0\, \theta(x_0) = \mathsf{D}_e \in T$, by (5)

Therefore, as by (a) T is consistent

$$\text{(B)} \quad \forall x_0\, \theta(x_0) = \mathsf{D}_e \notin T \text{ and } \neg \forall x_0\, \theta(x_0) = \mathsf{D}_e \notin T.$$

Discussion According to (A) the sentence $\forall x_0\, \theta(x_0) = \mathsf{D}_{\mathrm{g}(\Phi)}$ means intuitively: *the sentence named* $\mathsf{D}_{\mathrm{g}(\Phi)}$ *is true*, i.e., Φ *is true*. So D_e is the name of *the sentence named* $\delta(\mathsf{D}_n)$ *is not true*. This is equivalent to *the sentence named* D_e *is not true*, because $\delta(\mathsf{D}_n) = \mathsf{D}_e$ is true by (5). So the sentence named D_e affirms that D_e itself is not true, as is the case by the *liar* in the form of Eubolides (see Kleene [6], p. 39). So the sentence D_e is true iff D_e is not true, (6). To avoid a contradiction it and its negation must therefore be excluded from the theory T, i.e., (B) follows.

REFERENCES

[1] Germano, G., *Untersuchungen sur axiomatischen Präzisierung des Begriffs der Grundkenntnis in der Mathematik*, Dissertation, Münster i.W. (1966).

[2] Germano, G., "Metamathematische Begriffe in Standardtheorien," *Archiv für Mathematische Logik und Grundlagenforschung*, vol. 13 (1970), pp. 22-38.

[3] Germano, G., "Incompleteness and truth definitions," *Theoria*, vol. 37 (1971), pp. 86-90.

[4] Germano, G., "A very short proof of a strong incompleteness theorem," *Notices of the American Mathematical Society*, vol. 17 (1970), pp. 833-834.

[5] Grzegorczyk, A., *Some classes of recursive functions*, Rozprawy Matematyczne IV, Warszawa (1953).

[6] Kleene, S. C., *Introduction to Metamathematics*, Amsterdam (1967).

[7] Mostowski, A., *Sentences Undecidable in Formalized Arithmetic*, Amsterdam (1964).

[8] Myhill, J., "A system which can define its own truth," *Fundamenta Mathematicae*, vol. XXXVII (1950), pp. 190-192.

[9] Tarski, A., "Der Wahrheitsbegriff in den formalisierten Sprachen," *Studia Philosophica*, vol. 1 (1936), pp. 261-403.

[10] Tarski, A., in collaboration with A. Mostowski and R. M. Robinson, *Undecidable Theories*, Amsterdam (1953).

Laboratorio di Cibernetica
Consiglio Nazionale delle Ricerche
Arco Felice, Napoli, Italy