CAMBRIDGE

Cambridge University Press
978-1-107-19765-7 — The Epistemic Lightness of Truth: Deflationism and its Logic
Cezary Cieśliński
Excerpt
More Information

# 1 Preliminaries

This preliminary chapter introduces the notation and basic terminology. Apart from this, I will also formulate (usually without proofs) some classical results, which will be referred to in this book. It should be emphasised at the start that this is not intended to be a comprehensive overview of any particular discipline or area of research. As a matter of fact, the main and often the only criterion motivating the choice of the material is whether a given concept (or a lemma, or a theorem) will be useful in the chapters to follow.

## 1.1 *Peano Arithmetic*

The first definition describes the language of first-order arithmetic; in the next move, a concrete arithmetical theory will be characterised: Peano arithmetic.

DEFINITION 1.1.1. The language of first-order arithmetic, denoted here as $L_{PA}$, contains the usual logical vocabulary (quantifiers, connectives, brackets, and variables $v_0, v_1 \ldots$). The set of primitive extralogical symbols of $L_{PA}$ is defined as $\{`+\text{'}, `\times\text{'}, `0\text{'}, `S\text{'}\}$; in effect, it contains symbols for addition, multiplication, zero, and the successor function, respectively.

Terms, formulas and sentences of $L_{PA}$ are defined in the usual style (in particular, sentences of $L_{PA}$ are defined as formulas without any free variables). The expressions $Var$, $Tm$, $Tm^c$, and $Sent_{L_{PA}}$ will be used as referring (respectively) to the sets of variables, terms, constant terms, and sentences of $L_{PA}$. In general, for a theory $Th$, the expressions $L_{Th}$ and $Sent_{L_{Th}}$ will refer to the language of $Th$ and to the set of sentences of the language of $Th$.

The next definition introduces Peano arithmetic.

DEFINITION 1.1.2. Peano arithmetic ($PA$) is defined as the theory with the following arithmetical axioms:[1]

1 Apart from that, the set of axioms of $PA$ will contain the axioms of first-order logic.

1. $\forall x\ S(x) \neq 0$
2. $\forall x,y\ [S(x) = S(y) \rightarrow x = y]$
3. $\forall x\ x + 0 = x$
4. $\forall x,y\ x + S(y) = S(x + y)$
5. $\forall x\ x \times 0 = 0$
6. $\forall x,y\ x \times S(y) = (x \times y) + x$
7. $\{[\varphi(0) \wedge \forall x\ (\varphi(x) \rightarrow \varphi(S(x)))] \rightarrow \forall x\ \varphi(x) : \varphi(x) \in L_{PA}\}$

The last item is the set of arithmetical sentences falling under the schema of mathematical induction. Since there are infinitely many such sentences, the axiomatisation given here is patently not finite.[2]

The language of first-order arithmetic, as characterised in Definition 1.1.1, does not contain any numerals except for the symbol '0' (that is, it does not contain terms '1', '2' etc.). However, the notion of a numeral – a canonical term denoting a number – can be defined in the following way:

DEFINITION 1.1.3. A numeral is an arbitrary term of $L_{PA}$ of the form '$S \ldots S(0)$', i.e. a term obtained by preceding a symbol '0' with (arbitrarily many) successor symbols. If the number of successor symbols in a numeral equals $n$, the numeral will be abbreviated as $\overline{n}$.

Some schema of coding (or Gödel numbering) will be tacitly assumed throughout the book. It is possible to define a procedure, which starts with assigning numbers to primitive expressions of $L_{PA}$ and then extending the assignment to cover more complex syntactical objects. Eventually unique natural numbers become assigned to terms, formulas, and sequences of formulas (including proofs).[3] In effect it becomes possible to view some statements of first-order arithmetic as assertions about syntax.[4]

Truth predicate will be understood in this book as applying to syntactic objects, namely, to sentences.[5] Accordingly, a theory of syntax forms a

2 Moreover, in this respect the axiomatisation cannot be improved: it is known that Peano arithmetic is not finitely axiomatisable. See (Hájek and Pudlák 1993, p. 164), Corollary 2.24.

3 The classical method employs prime factorisation: a finite sequence of numbers $(n_1 \ldots n_k)$ will be coded by the number $2^{(n_1+1)} \times 3^{(n_2+1)} \times \ldots \times p_k^{(n_k+1)}$, with $p_k$ being the $k$-th prime.

4 I will not describe the details of coding here; they can be found, e.g. in (Kaye 1991).

5 Choosing sentences instead of propositions brings simplicity, although it should be admitted that this is not a philosophically innocent decision. In particular, Halbach (2011, p. 12) observes that the modal status of disquotation sentences (like "'Snow is white' is true if and only if snow is white") depends on whether truth is ascribed to a proposition or to a sentence, with some philosophers arguing that only with the first option the disquotation sentences become *necessary*.

necessary base for the theory of truth. Peano arithmetic is one of the theories suitable for this role, with the reason being that basic syntactic properties and relations are recursive, and Peano arithmetic is strong enough to represent them. The exact definition of the notion of a recursive set will not be given here; let me emphasise only that, in intuitive terms, a set is recursive if there is an algorithm which decides, for an arbitrary number $n$, whether or not $n$ belongs to this set. In what follows, I will describe only the important notion of representability together with its basic properties, treating the concept of a recursive set as given.

DEFINITION 1.1.4. A set of natural numbers $Z$ is representable in an arithmetical theory $Th$ iff there is a formula $\varphi(x)$ of the language of $Th$, with one free variable, such that for every natural number $n$:

1. if $n \in Z$, then $Th \vdash \varphi(\overline{n})$,
2. if $n \notin Z$, then $Th \vdash \neg\varphi(\overline{n})$.

With these conditions satisfied, we say also that $\varphi(x)$ represents $Z$ in $Th$.

Before formulating the representability theorem, let me introduce the familiar arithmetical hierarchy.

DEFINITION 1.1.5 (Arithmetical hierarchy).

- A bounded quantifier is a quantifier of the form '$Qx < y$', for $Q \in \{\forall, \exists\}$.
- A formula $\varphi$ belongs to the class $\Delta_0$ iff all the quantifiers in $\varphi$ are bounded. (We stipulate also that, by definition, $\Delta_0 = \Sigma_0 = \Pi_0$.)
- A formula $\varphi$ belongs to the class $\Sigma_{n+1}$ iff for some $\psi \in \Pi_n$ and for some sequence of variables $a$, $\varphi$ has a form '$\exists a \psi$'.
- A formula $\varphi$ belongs to the class $\Pi_{n+1}$ iff for some $\psi \in \Sigma_n$ and for some sequence of variables $a$, $\varphi$ has a form '$\forall a \psi$'.

$\Sigma_n$ and $\Pi_n$ classes were characterised here as containing only formulas of a rather special syntactic type. Observe in particular that Definition 1.1.5 does not introduce any closure of these classes under provable equivalence, and for this reason $\Sigma_n$ and $\Pi_n$ classes do not exhaust the set of all formulas (clearly there exist formulas whose syntactic form is altogether different, for example '$\exists x\, x = x \wedge \exists x\, x = x$' is neither $\Sigma_n$ nor $\Pi_n$). Nevertheless, it is possible to show that every formula is provably (in $PA$) equivalent to some $\Sigma_n$ (or $\Pi_n$) formula.

The following theorem is crucial for appreciating Peano arithmetic's role as a theory of syntax.

THEOREM 1.1.6 (Representability of recursive sets). For every recursive set $X$ of natural numbers, there is a $\Sigma_1$ formula representing $X$ in PA.[6]

Since a lot of basic syntactic properties are recursive, this gives us the means to build a theory of syntax inside $PA$. In particular, the following properties and relations are recursive:

- $x$ is a negation of $y$,
- $x$ is a conjunction of $y$ and $z$,
- $x$ is a variable, $x$ is a term, $x$ is a formula,
- $x$ is a numeral denoting a number $y$,
- $x$ is the result of substituting a term $t$ for a variable $v$ in a formula $z$.

Accordingly, Theorem 1.1.6 guarantees the existence of arithmetical formulas representing these syntactical properties and relations (they will be denoted, respectively, as $x = neg(y)$, $x = Conj(y,z)$, $Var(x)$, $Tm(x)$, $Fm(x)$, $x = name(y)$, and $x = sub(z,v,t)$). The road is open to building a theory of syntax inside $PA$.

The following application of the representability theorem will be of particular importance.

DEFINITION 1.1.7. Given a fixed recursive set $Ax(Th)$ axiomatising a theory $Th$, '$Prov_{Th}(x,y)$' is a formula of the language of $PA$ which represents in $PA$ the recursive relation '$d$ is a proof of $\varphi$ from $Ax(Th)$'. Given a formula '$Prov_{Th}(x,y)$', '$Pr_{Th}(y)$' is defined as the formula '$\exists x Prov_{Th}(x,y)$'.[7]

It should be stressed that by this definition, '$Prov_{Th}(x,y)$' is just *any* formula representing the relation of being a proof. For a given axiomatisation of $Th$, there will be many such formulas, sometimes with importantly different properties. The same concerns the provability formulas '$Pr_{Th}(y)$' – it is often important to keep in mind that it is not a uniquely determined single expression of $L_{PA}$.

In what follows I am not going to distinguish between formulas and their Gödel numbers (for all practical aims, I will just assume that formulas *are* Gödel numbers). Sometimes in this book square corners will be used for

6 For the proof, see (Kaye 1991, pp. 36–37).

7 Strictly speaking, for two different axiomatisations $Ax_1(Th)$ and $Ax_2(Th)$ of one and the same theory $Th$ we would need two different formulas '$Prov_{Ax_1(Th)}(x,y)$' and '$Prov_{Ax_2(Th)}(x,y)$', representing the relations of being a proof from the respective sets of axioms. I skip here this complication, noting only that the notation '$Prov_{Th}(x,y)$' presupposes a concrete, fixed axiomatisation of $Th$.

numerals denoting syntactic objects. Thus, if $\varphi$ is a formula, the notation $\ulcorner\varphi\urcorner$ is reserved for a numeral denoting $\varphi$. In addition, Feferman's dot notation will be occasionally employed. Thus, let $\varphi(x)$ and $\psi(x)$ be formulas. The expression:

$$\varphi(\ulcorner\psi(\dot{x})\urcorner)$$

will be treated as an abbreviation of

$$\exists y,z[y = name(x) \wedge z = sub(\ulcorner\psi(x)\urcorner, x, y) \wedge \varphi(z)].$$[8]

In some contexts, what is needed is not an arbitrary provability formula (build over an arbitrary proof predicate), but a predicate with some special properties. In such cases this will be stipulated explicitly. Some important constraints are listed in the next definition.

DEFINITION 1.1.8 (Derivability conditions). Given an axiomatisable theory $Th$ (in the language $L_{Th}$) extending $PA$, the following three statements will be called 'derivability conditions' for the predicate '$Pr_{Th}(x)$':

$(D_1)$ For every $\psi \in L_{Th}$, if $Th \vdash \psi$, then $PA \vdash Pr_{Th}(\ulcorner\psi\urcorner)$,
$(D_2)$ $\forall\psi,\varphi \in L_{Th}\ PA \vdash (Pr_{Th}(\ulcorner\varphi \to \psi\urcorner) \wedge Pr_{Th}(\ulcorner\varphi\urcorner)) \to Pr_{Th}(\ulcorner\psi\urcorner)$,
$(D_3)$ $\forall\varphi \in L_{Th}\ PA \vdash Pr_{Th}(\ulcorner\varphi\urcorner) \to Pr_{Th}(\ulcorner Pr_{Th}(\varphi)\urcorner)$.

Any provability predicate $Pr_{Th}(x)$ satisfying all three derivability conditions will be called 'standard'.

It is possible to show that the 'natural' provability predicate, defined in $PA$ in a way which closely mimics the usual, external definition of provability, is standard.[9]

8 Informally, this could be expressed as '$\varphi$ is true about the (Gödel number of the) result of substituting a numeral denoting $x$ for a free variable in $\psi$'. Observe that, in effect, the expression '$\varphi(\ulcorner\psi(\dot{x})\urcorner)$' contains $x$ as a free variable. If we used '$\varphi(\ulcorner\psi(x)\urcorner)$' instead, we would not obtain the same effect, as '$\ulcorner\psi(x)\urcorner$' is just a numeral – a constant term without any free variable inside.

9 For such a predicate, the basic formula '$Prov_{Th}(x,y)$' can be defined as stating (roughly): '$x$ is a finite sequence such that every element of $x$ is either an axiom of $Th$ or a logical axiom or it can be obtained from earlier elements of the sequence by a given rule of inference'.

The following lemma is crucial in many applications.

LEMMA 1.1.9 (Diagonal lemma). Let $Th$ be an extension of $PA$ (possibly in a richer language). For every formula $\varphi(x)$ of the language of $Th$, there is a sentence $\psi$ of the language of $Th$ such that:

$$Th \vdash \psi \equiv \varphi(\ulcorner \psi \urcorner).^{10}$$

It should be stressed that the formulation given here covers also cases in which the theory in question is formulated in a language richer than that of first-order arithmetic. In particular, the possibility of applying the diagonal lemma to truth theories (in the language with the truth predicate) will be important to us. It is worth mentioning that in such a case the theory needed to prove the biconditional '$\psi \equiv \varphi(\ulcorner \psi \urcorner)$' is a very weak extension of $PA$, obtained by adding to the axioms of $PA$ just the logical axioms in the extended language.

The diagonal lemma is employed in typical proofs of two famous incompleteness theorems, which are formulated below.

THEOREM 1.1.10 (Gödel-Rosser first incompleteness theorem). Let $Th$ be a consistent, axiomatisable extension of $PA$. Then there is a sentence $\psi \in L_{PA}$ such that neither $\psi$ nor its negation is provable in $Th$.

The theorem gives the information that no axiomatisable, consistent extension of Peano arithmetic will decide all arithmetical sentences. The sentence $\psi$, independent from $Th$, is obtained by diagonalising Rosser's provability predicate. Given a provability predicate $Prov_{Th}(x,y)$, define:

$$Prov^R_{Th}(x,y) =_{def} Prov_{Th}(x,y) \wedge \forall z < x \neg Prov_{Th}(z, \neg y).$$

Rosser's provability predicate can be defined by the condition:

$$Pr^R_{Th}(y) =_{def} \exists x Prov^R_{Th}(x,y).$$

It turns out that a sentence $\psi$ provably (in $Th$) equivalent to $\neg Pr^R_{Th}(\ulcorner \psi \urcorner)$ will be independent of $Th$.

A somewhat weaker result is obtained by diagonalising on an arbitrary predicate $Pr_{Th}(x)$ from Definition 1.1.7. It is known that any sentence $\psi$ provably equivalent to $\neg Pr_{Th}(\ulcorner \psi \urcorner)$ is not provable in $Th$ if only $Th$ is consistent; however, the negation of such a $\psi$ might be provable if $Th$ is $\omega$-inconsistent.[11] The meaning of this last notion is explained in what follows.

10 For more details and the proof, see (Hájek and Pudlák 1993, p. 158ff).
11 For details, the reader is referred to (Smoryński 1977).

DEFINITION 1.1.11. A theory $Th$ containing $PA$ is $\omega$-consistent iff for every formula $\varphi(x)$ of the language of $Th$:

$$\text{if for every natural number } n,\ Th \vdash \varphi(\overline{n}),\ \text{then } Th \nvdash \exists x \neg\varphi(x).$$

As it happens, $\omega$-inconsistency of a theory does not imply that the theory in question is inconsistent. However, the basic problem with $\omega$-inconsistent theories is that even if consistent, they admit no standard interpretation – they cannot be interpreted in the standard model of arithmetic (see Observation 1.2.4).

In this book the name 'Gödel sentence' will be reserved for an arbitrary $G$ satisfying the following condition.

DEFINITION 1.1.12. Let $Th$ be an axiomatisable extension of $PA$. A Gödel sentence for $Th$ will be an arbitrary sentence $G$ such that

$$Th \vdash G \equiv \neg Pr_{Th}(\ulcorner G \urcorner).$$

Gödel's second incompleteness theorem concerns the unprovability of consistency. The formulation is given next.

THEOREM 1.1.13 (Gödel's second incompleteness theorem). Let $Th$ be any axiomatisable, consistent extension of $PA$. Let $Pr_{Th}(x)$ be a standard provability predicate for $Th$ (under a chosen recursive axiomatisation of $Th$). Denote as '$Con_{Th}$' the sentence '$\neg Pr_{Th}(\ulcorner 0 = 1 \urcorner)$'. Then $Th \nvdash Con_{Th}$.

Given that the derivability conditions (see Definition 1.1.8) are satisfied, the choice of '$0 = 1$' for the characterisation of the sentence '$Con_{Th}$' is not important, and any contradiction would be just as suitable.[12] The restriction to standard provability predicates (satisfying derivability conditions) in the formulation of the theorem is important. On the one hand, if the provability predicate is standard, then $Con_{Th}$ will be equivalent (provably in $Th$) to an arbitrary Gödel sentence for $Th$, and since the latter is not provable in a consistent theory $Th$, the same holds for $Con_{Th}$. On the other hand, without such a restriction counterexamples to Theorem 1.1.13 could be given. It is known, for example, that if we take $Pr^{R}_{Th}(x)$ as our starting point and define '$Con^{R}_{Th}$' as the sentence '$\neg Pr^{R}_{Th}(\ulcorner 0 = 1 \urcorner)$', then $Th \vdash Con^{R}_{Th}$.[13]

12 For an arbitrary sentence $\varphi$ disprovable in $Th$, we have: $PA \vdash Pr_{Th}(\ulcorner \varphi \urcorner) \equiv Pr_{Th}(\ulcorner 0 = 1 \urcorner)$.

13 For more details about the second incompleteness theorem, see, e.g. (Boolos et al. 2002, p. 247ff); see also (Cieśliński 2002) and (Cieśliński and Urbaniak 2013). For the provability of Rosser consistency, see, e.g. (Smoryński 1977, p. 841).

From the incompleteness phenomena we move now to *completeness*. The next two theorems characterise an important completeness property of arithmetical theories.

THEOREM 1.1.14 ($\Sigma_1$-completeness). Every $\Sigma_1$ sentence true in the standard model of arithmetic is provable in Peano arithmetic.

For the proof, see (Rautenberg 2006, p. 186).[14] In addition, it turns out that Theorem 1.1.14 can be formalised in $PA$.

THEOREM 1.1.15 (Formalised $\Sigma_1$-completeness). There is a standard provability predicate $Pr_{PA}(x)$ such that for every $\Sigma_1$ sentence $\psi \in L_{PA}$, $PA \vdash \psi \rightarrow Pr_{PA}(\psi)$.

For details the reader is referred to Section 7.1 of (Rautenberg 2006)[15] – one of the few textbooks giving a detailed proof of the derivability conditions and formalised $\Sigma_1$-completeness of Peano arithmetic.

Let us end this section with another useful classical theorem where the assumption of the standardness of the provability predicate is essential again.

THEOREM 1.1.16 (Löb's theorem). Let $Th$ be an axiomatisable, consistent extension of $PA$ and let $Pr_{Th}(x)$ be a standard provability predicate. Then for every formula $\beta$ of the language of $Th$:

$$Th \vdash Pr_{Th}(\ulcorner \beta \urcorner) \rightarrow \beta \text{ iff } Th \vdash \beta.$$[16]

## *1.2 Model Theory*

The reader is assumed to be familiar with the concept of a mathematical structure and with the notion of truth in a model. In this book I will not use separate symbols for models and their universes. In particular, the symbol $N$ will be employed as referring to the standard model of arithmetic but also to the set of natural numbers.

Two definitions given in what follows introduce some basic terminology. A signature (or a type) of a given mathematical structure is the information about the number and the arity of the relations, the operations and the constant elements of the structure.[17] Signatures can be assigned also to

14 Theorem 3.1 in Rautenberg's book is even stronger than that: it attributes $\Sigma_1$ completeness to Robinson's arithmetic, which is a finitely axiomatisable subtheory of $PA$.

15 See especially Theorem 1.2 on p. 215.

16 For the proof see (Boolos et al. 2002, p. 237); see also (Cieśliński 2003) for a discussion of Löb's theorem in set theory.

17 For a full definition, see (Adamowicz and Zbierski 2011, pp. 11–12).

languages and if a given language $L$ has the same signature as a mathematical structure $S$, we say that $S$ is a model of $L$.

DEFINITION 1.2.1. A set $X$ is definable with parameters in a model $M$ of the language $L$ iff there is a formula $\varphi(x, y_1 \dots y_k) \in L$ and $a_1 \dots a_k \in M$ such that $X = \{z : M \models \varphi(z, a_1 \dots a_k)\}$.

DEFINITION 1.2.2. Let $M$ be a structure with the same signature as a given first-order language $L$. We define:

- $Th(M) = \{\psi \in L : M \models \psi\}$. The set $Th(M)$ is called the theory of $M$.
- $L(M)$ – the language of $M$ – is an extension of $L$ with a set of new constants, corresponding to all elements of $M$. (In effect, we enrich $L$ with the set of constants $\{c_a : a \in M\}$.)
- $ElDiag(M)$ – the elementary diagram of $M$ – is defined as the set $\{\psi \in L(M) : M \models \psi\}$.[18]

The next definition characterises the notions of an extension and an expansion of a model. Roughly, extensions add new elements; expansions leave the old model intact, adding only interpretations of new symbols in the old model.

DEFINITION 1.2.3.

- A model $M$ is an extension of a model $K$ (or: $K$ is a submodel of $M$) iff the universe of $K$ is a subset of the universe of $M$ and the relations and functions of $K$ are just relations and functions of $M$ restricted to the universe of $K$.
- A model $M$ is an expansion of a model $K$ iff the only difference between $M$ and $K$ is that $M$ contains new relations, functions or constant elements, absent in $K$.

Truth-expansions of models of $PA$ will be particularly important. Given a model $(M, +_M, \times_M, S_M, 0_M)$ of $PA$, I will abbreviate as $(M, T)$ the expansion $(M, +_M, \times_M, S_M, 0_M, T_M)$ of the initial model. In such a context $T$ will be a subset of $M$ which serves as an interpretation of the truth predicate.

Definition 1.1.11 introduced the notion of an $\omega$-consistent theory. We noticed that $\omega$-inconsistency does not imply inconsistency: if $\omega$-inconsistent theories are not attractive, it is not because they are inconsistent. The reason

18 The definition of $ElDiag(M)$ resembles that of $Th(M)$; the only difference lies in taking into account all sentences of $L(M)$ instead of $L$.

to be dissatisfied with $\omega$-inconsistent theories is given in the observation that follows.

OBSERVATION 1.2.4. If $Th$ is $\omega$-inconsistent, then the standard model of arithmetic cannot be expanded to a model of $Th$.

PROOF. Assume that for all $n \in N$, $Th \vdash \varphi(\overline{n})$ but $Th \vdash \exists x \neg\varphi(x)$. Let $N^*$ be an expansion of $N$ such that $N^* \models Th$. Pick an $a$ such that $N^* \models \neg\varphi(a)$. Then $a \in N$ (since $N^*$ is an expansion of $N$), but this is impossible, because then by assumption $N^* \models \varphi(a)$. ⊣

Since the standard model of arithmetic is typically meant to provide the intended interpretation for theories extending $PA$, the lack of such an interpretation is a quite undesirable trait.

Later on I will sometimes make use of the soundness properties of $PA$ and its extensions. In general, soundness of a theory means that theoremhood implies truth or validity. Here the emphasis will be mostly on truth of arithmetical sentences in the standard model. The definition that follows introduces the notion of soundness with respect to a given class of sentences.

DEFINITION 1.2.5. Let $\Gamma$ be a class of arithmetical sentences. A theory $Th$ is $\Gamma$-sound iff for every arithmetical sentence $\psi$ belonging to $\Gamma$, if $Th \vdash \psi$, then $\psi$ is true in the standard model of arithmetic.

A discussion of sets, even infinite ones, can be sometimes carried out in an arithmetical language inside a given (nonstandard) model of Peano arithmetic. Let '$y = p_x$' be an arithmetical formula with the meaning '$y$ is the $x$th prime number'; abbreviate as '$x|y$' the arithmetical formula '$x$ divides $y$'. Then we define:

DEFINITION 1.2.6. For every $M$, for every $a \in M$, for every set of natural numbers $Z$, $a$ codes $Z$ in $M$ iff

$$Z = \{n : M \models p_n | a\}.$$

Instead of '$p_x|a$' I will usually write: '$x \in a$', treating the latter formula as belonging to the language of arithmetic.

This idea of coding permits to reproduce some set theory inside models of arithmetic. Observe that in the standard model of arithmetic, only finite