Interpretation and Truth in Set Theory

Rodrigo A. Freire

Abstract The present paper is concerned with the presumed concrete or interpreted character of some axiom systems, notably axiom systems for usual set theory. A presentation of a concrete axiom system (set theory, for example) is accompanied with a conceptual component which, presumably, delimitates the subject matter of the system. In this paper, concrete axiom systems are understood in terms of a double-layer schema, containing the conceptual component as well as the deductive component, corresponding to the first layer and to the second layer, respectively. The conceptual component is identified with a criterion given by directive principles. Two lists of directive principles for set theory are given, and the two double-layer pictures of set theory that emerged from these lists are analyzed. Particular attention is paid to set-theoretic truth and the fixation of truth-values in each double-layer picture. The semantic commitments of both proposals are also compared, and distinguished from the usual notion of ontological commitment, which does not apply. The approach presented here to the problem of concrete axiom systems can be applied to other mathematical theories with interesting results. The case of elementary arithmetic is mentioned in passing.

1 Introduction

A foundational analysis of meaning and truth in axiomatic theories usually begins with a division of axiom systems in two groups: An axiom system is said to be *concrete* if its language is supposed to be interpreted in a specific way. In opposition, an axiom system is said to be *abstract* if it is not supposed to be interpreted in a specific way. For example, when explaining the construction of axiom systems in mathematics, Shoenfield writes:

> We have so far supposed that we have definite concepts in mind. Even so, it may be possible to discover other concepts which make the axioms true. In this case, all the theorems proved will also be true for these new concepts. This has led mathematicians to frame axiom systems

R. A. Freire (✉)
Philosophy Department, University of Brasília, Brasília, Brazil
e-mail: rodrigofreire@unb.br

W. Carnielli and J. Malinowski (eds.), *Contradictions, from Consistency to Inconsistency*, Trends in Logic 47, https://doi.org/10.1007/978-3-319-98797-2_9

> in which the axioms are true for a large number of concepts. A typical example is the set of axioms for a group. We call such axiom systems *modern* axiom systems, as opposed to the *classical* axiom systems discussed above. Of course, the difference is not really in the axiom system, but in the intentions of the framer of the systems. ([12], p. 2)[1]

I agree with Shoenfield that it is unproblematic that the difference is not really in the systems. However, how could one fix the system's intended interpretation? There are obvious problems with the claim that the intentions of the framer of the system already fix the interpretation. For it is not enough to intend. It is not clear how intentions could fix an interpretation, for the intention to present an axiom system about some specific subject matter may not suffice to interpret the axiom system.

When we are introduced to axioms for (first-order) arithmetic we are (usually) supposed to learn that they are about definite concepts of sum and product of natural numbers. But, using Shoenfield's formulation, there are other concepts which makes the usual axioms of arithmetic true. An axiom system is constituted by sentences which are purely linguistic objects, and its interpretation, what it is intended to be about, is therefore a matter of language convention. The acquired meaning of linguistic symbols is given by some kind of convention. What is the language convention according to which it is possible to interpret axioms for arithmetic in the usually intended way? One such language convention must *fulfill* intentions, but it cannot be *made* of intentions. For it seems clear that we do not learn intentions, whatever this could mean, when we learn that an axiom system for arithmetic talks about the definite concepts of sum and product of natural numbers, because it is not clear how to constitute a language convention from intentions.

This paper is first of all concerned with the following

Problem 1: *How could we explicate the commitments latent in the interpretation of an axiom system for set theory we happen to have inherited?*

If an axiom system is supposed to talk about some definite concepts, then there must be something prior to the system in question that we have to adopt in order to interpret its language. Otherwise, how could the sentences of an axiom system acquire the meaning that fulfills the intention to talk about specific concepts? As I have argued above, although this prior standard that we have adopted when we learn to interpret an axiom system for set theory must fulfill the relevant intentions, it cannot be a plurality of intentions. Now, the relevant question is: What should we adopt in order to interpret an axiom system for set theory in some specific way?

Of course, one could say "why bother? If we do not know how to make sense of the notion of interpreted formal system, just forget about it. It does not make any difference when we are doing the technical work with formal systems." Unfortunately, the problem of making sense of the notion of interpreted formal system is directly related to the problem of fixing truth-values. Based on Gödel's incompleteness theorems, arithmetical truth is generally held to go beyond provability, implying that the axiom systems for arithmetic must be regarded as concrete. Therefore, there must be an interpretation of the language of arithmetic fixing truth-values beyond provability.

[1]The terms *concrete* and *abstract* are used, for example, by Tait in [13], p. 90.

Making sense of the notion of interpreted formal system is a required component in any explanation of this transcendent nature of arithmetical truth.

Notice that problem 1 is only asking how can we understand the notion of concrete axiom system for set theory, that is, how can we explicate interpreted character of the language of set theory that excludes from consideration at least some concept of set membership which nevertheless makes the axioms true. The thesis that there is one correct concrete axiom system for set theory is a substantive thesis that I will discuss later.

In general, we say that a *sentence* is true with respect to a *prior standard* if an *agreement* between sentence and standard obtains. Otherwise, we say that the sentence is false with respect to the instituted standard. The commitments latent in an interpretation of an axiom system for set theory amounts to such a prior standard.

There is a traditional position on this problem according to which an interpreted axiom system for set theory is an axiom system endowed with a model, or a class of models which constitute a prior standard for the correctness of the axiom system, existing independently of our mathematical practice. These *independent models* cannot be the usual mathematical models. In fact, usual mathematical models are mathematical objects living inside set theory, and cannot be used to define the presumed concrete character of set theory. In this picture, the interpretation of an axiom system for set theory is committed to the existence of independent models.

Furthermore, in this framework, we are assumed to have some kind of "nonpropositional grasp", or "mathematical intuition" of independent models, from which we interpret the axiom systems and ground the axioms. This view will be discussed later. For now, it is enough to say that I think this position is inadequate. Models for set theory are very complicated, there is no simple representation of such a model, and it is not reasonable to say that models for set theory are, historically or conceptually, prior to the axioms. It is not clear how we could have some kind of "nonpropositional grasp" of those models, and how we could extract a language convention connecting the axiom system and its models from it.

I will work out a completely different solution to problem 1 based on the following thesis: If an axiom system is supposed to be interpreted in a specific way, then there must be a public *criterion*, given as principles which are considered unambiguous, objective and based on the standard practice of the relevant mathematical theory[2] and prescribe the interpretation of the system, that is, the intended class of models. What do we learn when we learn to interpret an axiom system for set theory is a list of principles that prescribe that interpretation. It is not necessary to assume the existence of standard models independent of our mathematical practice in order to *prescribe*, through a criterion, what would count as one.

In Sect. 2, two lists of principles corresponding to two criteria for interpreting the axiom systems for set theory will be explicitly stated, and these criteria constitute two

[2]The standard practice of set theory, that is the mathematical activity initiated by Cantor's seminal works, is *historically given*. The historical existence of this practice, considered as a merely historical phenomenon, is, of course, independent of the set-theoretic principles extracted from it. Therefore, there is no circularity in the relation between the set-theoretic principles that will be given here and the practice of set theory.

solutions for problem 1. The principles in those lists will be called *directive principles*. I assume, therefore, the priority of a list of directive principles over the corresponding class of standard models, and the axiomatization of a mathematical theory must be guided by such a list. The primary standard for the correctness of a mathematical theory is the corresponding list of directive principles, the corresponding class of standard models must be understood as a secondary standard for the correctness of the theory, and only insofar as it represents the principles. Each list of principles constitutes a direction to be pursued by the theory.

After stating explicitly the directive principles for the two lists, I will be concerned with the following:

Problem 2: *How and to what extent does the interpretation of the axiom systems for set theory committed to the criterion corresponding to given directive principles fix truth-values?*

In Sect. 3, an analysis of problem 2 will be given.

The analysis of problems 1 and 2 is independent of the substantive thesis according to which there is one correct interpretation of the axiom systems for set theory. However, it will be argued that the interpretation prescribed by each list of directive principles stated below is a plausible understanding of the axiom systems for set theory. Its plausibility ultimately comes from the fact that it is an interpretation prescribed by principles extracted from the *standard* practice of set theory. Therefore, the solutions to problem 1 given in this paper are at least historically correct in the sense that it is based on the thesis that the presumed concreteness (or classicality) of axiom systems for set theory must rest on instructions by which we understand sets according to Cantor, Dedekind, Fraenkel, Hilbert, Zermelo, etc. I defend the thesis that the organization of set theory is not ultimately based on, or committed to, a nonpropositional grasp of independent objects, nor on intentions, but on a criterion given as a list of set-theoretic principles extracted from the standard practice of this mathematical theory.

Vann McGee has written a paper on a related subject [10]. In his paper's introduction McGee says:

> The internal problem is this: the realist conception supposes that the meaning of mathematical terms is fixed with sufficient precision to ensure that each sentence has a determinate truth value. Now whatever meaning a linguistic expression has it possesses in virtue of the thoughts and practices of human beings. Not all meaning is thus dependent on human thought and action – the fact that a red sky in the morning means stormy weather isn't a matter of convention – but the fact that the numeral '7' refers to the fourth prime is a matter of how we have chosen to use a symbol. So there must be something we think, do, or say that fixes the intended meaning of mathematical terms. How are we able to do this? Mathematical objects aren't like Fido, whom you can hold by the collar while you give him a name. Nonetheless, there is something we think, do, or say that connects concrete speech acts with their abstract referents. Until we can give at least a rudimentary account of how this is done, the realist doctrine that mathematical sentences have determinate truth values will remain deeply mysterious. ([10], pp, 35–36)

Although closely related, McGee's internal problem of realism and problem 1 on axiom systems for set theory are, strictly speaking, different problems. They are conceptually different, and there are at least two important practical differences between McGee's problem and problem 1. First, McGee's problem presupposes that each sentence has a rigid truth-value. This is not required for an axiom system to be concrete. In fact, an axiom system is not to be considered concrete if it is understood as talking about whatever structure satisfying the axioms; otherwise it is interpreted according to a criterion. Thus, to say that an axiom system is interpreted in a specific way is not to say that its subject matter is a unique structure up to isomorphism, or a class of structures which are all elementary equivalent, but that it is not the class of *all* structures which satisfy the axioms. Secondly, for both McGee's problem and problem 1, the concrete character of an axiom system must rest on mathematical practice, but for a criterion to be considered a solution to problem 1 it must prescribe a class of structures and be extracted from the *standard* practice of set theory. The latter condition is not required for a solution of McGee's problem. McGee's proposal is stated in general outline in the following paragraph:

> Knowing how to use mathematical terms in practical problem solving is an important component of our understanding of mathematical vocabulary, but it doesn't take us far enough. Something more is needed. The format of our proposed answer is this: What we learn when we learn mathematical vocabulary, apart from learning how to count and measure, is a body of mathematical theory. What else could the answer be? The meaning given to a mathematical term is wholly dependent upon our use of the term (unlike a term like "Fido", whose meaning depends partly on our usage and partly on causal connections beyond our control), and our practical uses of the term aren't enough to determine the truth values; so what else is left but our use of the term in theorizing? ([10], p. 40)

There are important similarities and dissimilarities between McGee's solution to his problem and the solutions to problem 1 that I will present in this paper, and the comparison of these solutions is illuminating. I am arguing that the solution to problem 1 is that a concrete axiom system is a formal system accompanied by a criterion given as a list of directive principles which gives the commitments latent in the interpretation of set theory and, although these principles must be given as set-theoretic principles extracted from standard practice of set theory, they are *not* a separate mathematical theory as in McGee's proposal. Another point is that McGee's solution is not faithful with respect to the standard practice of set theory. Indeed, McGee appeals to an *Urelement Set Axiom* ([10], p. 52), and this axiom is not to be found in the standard practice of set theory. On the other hand, the solution to problem 1 that will be presented shortly is not a solution to McGee's problem, because it does not fix all truth values. There would be more interesting things to say in this regard, but I shall not be occupied with an exegesis of McGee's paper.

2 Two Double-Layer Pictures for Set Theory

As I have already said in Sect. 1, I propose that a criterion, given by set-theoretic principles, must be the basis for a solution of problem 1, and these principles will be called directive principles. From now on, for the sake of definiteness, I will identify axiom systems with formal first-order systems, in the usual sense. The solution that I propose is roughly as follows: Formal systems for set theory can be considered as concrete if and only if they are, tacitly or not, accompanied by a criterion, given as a list of directive principles, which prescribe its interpretation, that is, the appropriate class of structures. There is a list of minimal *desiderata* that a proposed solution to problem 1 has to meet in order to be treated as a plausible understanding of the axiom systems for set theory: (i) A solution to problem 1 is based on set-theoretic principles, the directive principles, that must be explicitly stated. (ii) All these principles must be extracted from the standard mathematical practice of set theory, according to the thesis that, whatever meaning a linguistic expression in this theory has, it has in virtue of the standard practice of set theory. (iii) It is desirable that the instructions contained in the directive principles solving problem 1 relate to the formal systems for set theory in an unproblematic way, without appeal to any intuition of independent objects. (iv) One such solution must fix truth-values in a satisfactory way, that is, at least all arithmetical statements must have rigid truth-values in the interpretation given by the corresponding criterion.

Of course, since the directive principles are stated explicitly below, it follows that this solution meets *desideratum* (i). The following directive principles highlight the *production* of sets, which is undoubtedly a central aspect of set theory.

2.1 The First List of Directive Principles

1. A set is determined by its elements, which are sets themselves, and there is no infinite regress in this transitive determination.
2. An arbitrary choice of elements of a set determines a set, which is a subset of the original set.
3. An arbitrary replacement of each element of a set by a set determines a set.
4. All the elements of the elements of a set determine a set.
5. All the subsets of a set determine a set.
6. There is an infinite set.

The first thing to say about the criterion given by the directive principles above is that they are *not* an independent formal system for set theory, nor a natural language formulation of the axioms of a formal system. They are not even a separate theory but just give a criterion for interpreting the formal systems for set theory. This criterion *is not defining sets*, in any significant sense. Recall that, for the language of set theory to be understood as being about something more specific than whatever concept of set membership which makes the axioms true, there must be some prior standard, a criterion excluding some of these concepts, that we learn when we learn to interpret it. The directive principles can play the role of this prior standard: They are the basic directives for an axiomatic theory of sets that can be found in the main line of thought developed throughout set theory, and this is the only a priori justification needed for them, as they are not, in any philosophically significant sense, "absolute set-theoretic principles".[3]

Now, notice that the formal system ZFC can be obtained from the directive principles by formalization. For example, each instance of the replacement axiom is directly obtained from principle (3): Given a set A and a functional formula F, principle (3) gives the unique set B obtained from A by replacing each element by its F-image, which is the set required to exist by the replacement axiom. The only nontrivial case is the axiom of choice, but this axiom can also be obtained by a formalization of the principles: A choice set for a given set is a subset of the union of the set, and hence can be obtained by the use of directive principles (2) and (4). It can also be easily obtained from directive principle (3) alone, as the production of a choice set for a given set is the particular case of a replacement of each element of the given set by an element of itself.[4] The axiom of choice can be seen as an attempt at formalizing part of the *arbitrariness* present in directive principles (2) and (3). In Sect. 3 I will show that this arbitrariness cannot be fully formalized.

In order to make explicit the point that the criterion given by the directive principles above meets *desideratum* (ii), that is, that they are historically coherent with the formal systems for set theory, I will say a few words on the origins of the principles: Harward and Cantor, although *not working from an axiomatic perspective*, discussed what they considered to be very basic facts about sets. The former carried forward his discussion in an article at the Philosophical Magazine in 1905, and the latter in correspondence with Dedekind and Hilbert. The facts that they noted are very similar to the directive principles stated above: Both suggested something like principles (2) to (6). For example, Cantor in a letter to Dedekind ([2], p. 114) stated the following:

> Two equivalent multiplicities either are both "sets" or are both inconsistent.
> Every submultiplicity of a set is a set.
> Whenever we have a set of sets, the elements of these sets again form a set.

Directive principles (2), (3) and (4) are *collectively equivalent* to these three statements in Cantor's letter. Of course, principles (5) and (6) are also a basic part of Cantor's seminal works.

Principle (1) is the outcome of the guiding thought according to which the right way to think about a set is that it is, whatever its nature may be, a well-founded, extensional entity determined directly by its members. This thought can be reformulated as [A set is an object whose immediate constituents are its elements, and it is determined by its elements. Furthermore, since an object cannot be a constituent, immediate or

---

[3]This is an important point: The directive principles above give instructions for understanding formal sentences in usual set theory – these principles are *not* directly related to other set theories such as those theories of non-well-founded sets. Therefore, their application outside the framework of usual set theory is not justified.

[4]This is not accidental. In fact, with the exception of the empty set, which can be produced by directive principle (2) from any given set, all subsets produced by directive principle (2) from a given set can also be produced by directive principle (3). Indeed, it is easy to see that for each arbitrary choice of elements of a given set producing a nonempty set there is a replacement of the elements of the given set producing the same nonempty set.

not, of itself, it follows that sets should be extensional and well-founded objects.] The following famous passage of [3] is relevant for this thought and, consequently, for directive principle (1):

> By an "aggregate" (*Menge*) we are to understand any collection into a whole (*Zusammenfassung zu einen Ganzen*) M of definite and separate objects m of our intuition or our thought. These objects are called the "elements" of M. ([3], p. 85)

It seems clear from this passage that, according to Cantor, a set is determined by its elements.[5] The extensionality of sets is surely a component of the set theory developed by Cantor, but it seems that he nowhere explicitly talks about it. Dedekind states it as a fact about sets in his famous 'The Nature and Meaning of Numbers'. An explicit thought on the well-foundedness of sets is to be found on early drafts of Zermelo's axioms (preceding the publication of [14]) where, according to Moore ([11], p. 165), he had initially assumed, as a postulate, that a set could not be a member of itself.

I will not make any further comments on the origin of the directive principles in the given list. The above paragraphs are meant as a brief illustration of how these principles can be seen as the outcome of guiding thoughts on sets by the founding fathers of set theory, sometimes working in different perspectives, with different intentions and independently of each other. The directive principles are a primary expression of directions that led to the formal systems for set theory, and, as I understand it, one such expression must be part and parcel of this theory. This account for the possibility of interpretation of formal systems for set theory is *not* a theory about the psychological origin of the axioms of those formal systems. Accordingly, directive principles are not intentions, intuitions or dialectical considerations about sets; they constitute a *public* criterion for interpreting formal sentences in set theory and this historical note is intended to show their correctness with respect to the standard practice of this mathematical theory.

The codification of the directive principles in a formal (first-order) system give rise to nonlogical axioms in a formal language, but the principles *precede* their formal counterparts in the schema of set theory. The formal systems for set theory can be obtained from the directive principles by an operation of formalization. Since formalization does not appeal to any intuition of mathematical objects, the proposed solution meets *desideratum* (iii). Also, it must be obvious that we should not throw the directive principles away after formalizing set theory, unless we want to give up the possibility of interpreting the resulting formal systems according to those principles. Thus, I am proposing that set theory consists of two layers: The directive principles are the first layer, and the formal counterparts of the directive principles are the second layer. As it is already known, the formal system ZFC is one such formal

---

[5]It is also important to notice that in this passage Cantor seems to express the view according to which all sets are psychological objects. This is *irrelevant* for understanding a theory of sets: There may be concrete and abstract, psychological and physical sets, and a theory of sets must only be concerned with aspects common to all these possibilities. Therefore, the only important point extracted from this passage is that a set is, *whatever its specific nature may be*, determined by its elements without any further addition.

counterpart of the directive principles, which means that ZFC is in the second layer in this double-layer schema. Also, the arbitrariness present in directive principles (2) and (3) cannot be fully formalized, which means that the second layer in the double-layer schema is open-ended.

Furthermore, reinforcing what was said above, although the directive principles above can be equivalently written in the second-order language of set theory, they are *not* a kind of "second-order formal system". The directive principles do not constitute any kind of autonomous formal system in the sense that they do not play a deductive role in the double layer schema. Therefore, there is emphatically no attempt to replace the usual (first-order) formal systems for set theory by directive principles – this would be a very bad move because these principles cannot play the deductive role of a formal system. The double-layer schema of set theory stated above is based on the complementarity of directive principles and formal systems: The principles demarcate the subject matter of the formal systems for set theory by giving a criterion for excluding at least some concepts of set membership which nevertheless makes the axioms true, because they constitute stronger commitments, but they cannot replace their formal counterparts in proofs.

2.2 The Second List of Directive Principles

1. A set is determined by its elements, which are sets themselves, and there is no infinite regress in this transitive determination.
2. The elements of a set satisfying a property determines a set, which is a subset of the original set.
3. The replacement of each element of a set by a set given by a functional relation determines a set.
4. All the elements of the elements of a set determine a set.
5. All the subsets of a set determine a set.
6. Given a countable list of sets, there is a set determined by exactly those sets in the list.
7. Given a set there is a choice function on this set.

The general remarks about directive principles given above apply equally to this second list. Now, in contrast with the first list, sets given by replacement and separation are supposed to be *defined*. There is a stronger principle of infinity, principle (6), which can also be understood as an infinitary replacement axiom for countable sets, for it can be reformulated as follows: An arbitrary replacement of each element of a countable set by a set determines a set. One traditional argument for the replacement axiom states that without this axiom it is not possible to prove the existence of the set $\{\omega, \wp(\omega), \wp(\wp(\omega)), ...\}$. Of course, we can argue that there is nothing special with this countable set, and this is an evidence that the stronger principle of infinity is instituted in the practice of set theory, used to justify the replacement axiom. In fact, we can find the stronger principle of infinity expressed in [6], p. 46, footnote 5 (and also in [1]):

> A stronger axiom schema of infinity than **VI** is introduced in Fraenkel 27 (p. 114, Axiom **VII**c. Fraenkel's axiom is equivalent, on the basis of axioms **I**-**V**, to the schema which asserts, roughly, that every "denumerable collection of elements" is a set (Bernays 37–54 **III**).

Furthermore, the notion of countable list of sets is present in set theory from its beginnings. For example, Cohen says ([4], p. 64):

> It is perhaps not generally known, but Cantor's stimulus to study set theory arose from countable ordinals.

A choice principle concludes our list, which, of course, is extracted from the standard practice of the discipline.

A first-order (partial) formalization of the above principles results in ZFC, naturally. These principles can be equivalently written in the infinitary language $L_{\omega_1\omega_1}$. The resulting list of sentences is equivalent to ZFC plus

$$\forall x_0 \forall x_1 ... \exists y \forall z (z \in y \leftrightarrow z = x_1 \vee z = x_2 \vee ...).$$

This second list of principles has the virtue that it can be equivalently written in the language $L_{\omega_1\omega_1}$, while the first list requires the much more complex second-order language. In other words, the meaningfulness of the criterion given by the second list of principles is a cheap assumption when compared to the meaningfulness of the criterion given by the first list. Unfortunately, based on a mathematical analysis of the proposed lists of directive principles, it is arguable that only the first list captures the conception of *cumulative hierarchy of sets*, which seems to be the basic concept in terms of which set-theoretic truth is usually understood. The analysis of both lists that will be given in Sect. 3 clarifies the differences in a precise way.

The first layer in each of these double-layer pictures gives a criterion for the interpretation of formal systems for set theory; the second layer is always the domain of proof and formalization within the finitary range. The double-layer pictures of set theory that I have proposed have nothing to do with the two layers of mathematics in the picture of Platonism as diagnosed by Tait in the criticisms of Benacerraf and Dummett:

> Benacerraf and Dummett seem to me to be typical of those who adopt a particular picture of Platonism. The picture seems to be that mathematical practice takes place in an object language. But this practice needs to be explained. In other words, the object language has to be interpreted. The Platonist's way to interpret it is by Tarski's truth definition, which interprets it as being about a model – a Model-in-the-Sky – which somehow exists independently of our mathematical practice and serves to adjudicate its correctness. So there are two layers of mathematics: the layer of ordinary mathematical practice in which we prove propositions such as [There is a prime number greater than 10] and the layer of the Model at which [There is a prime number greater than 10] asserts the 'real existence' of a number. ([13], p. 67)

Later on, Tait concludes

> The myth of the Model tends to get attached to Platonism (or at least to 'epistemological' Platonism in the sense of Steiner (1975)) because the view that mathematics is about things

> like the system of numbers is compared with the view that propositions about sensible things are about the physical world; and here there is a tendency to believe that there *is* such a nonpropositional grasp, namely, sense perception, which does endow meaning on what we say and to which we appeal to determine truth. But I hope that, if not what I have said, then Wittgenstein's critique of this view of discourse about sensibles will convince the reader that it is inadequate. ([13], p. 74)

There is no such thing as the layer of the "Model(s)-in-the-Sky" in the double-layer accounts of set theory given here, and there is no appeal to a nonpropositional grasp of independent objects. First of all, the double-layer accounts given here are not committed to independent objects, but are only committed to the objectivity of the criteria given by directive principles. Furthermore, there is no need of any kind of nonpropositional grasp, which is usually needed in a schema in which the possibility of interpretation is alienated from formal systems. If the second layer of a double-layer schema consists of first-order counterparts of the first layer, as is the case here, then the formal system's corresponding interpretation is not alienated from them and there is no nonpropositional grasp involved.

With respect to the particular picture of Platonism which seems to be adopted by Benacerraf and Dummett in their criticisms, Tait writes

> ... Needless to say, it is not this version of Platonism that I am defending or that I even understand. Thus, I should not be understood to be taking part in any realism/antirealism dispute, since I do not understand the ground on which such disputes take place. As a mathematical statement, the assertion that numbers exist is a triviality. What does it mean to regard it as a statement outside of mathematics? ([13], p. 68)

The directive principles are not subject to Tait's question, because they are neither statements outside set theory, nor ways of stepping outside this mathematical framework. These principles just give us a criterion for the interpretation of formal systems for set theory. When mathematically analyzed, each list of directive principles gives a criterion separating a class of *standard* structures and giving the truth conditions for each sentence, as it will be shown in Sect. 3.

2.3 Other Views on Set Theory and Mathematical Truth

A possible alternative to a double-layer schema of set theory is, of course, a single-layer account of it, according to which set theory consists only of a layer of first-order formal systems, which can be either static or evolving in time, such that their formal languages are not supposed to be about something more specific than all possible models of their axioms. It seems to me that this single-layer picture is inadequate in several ways. I will mention three: First, it is historically problematic, because the original framers of set theory did not think about set theory that way, as an abstract axiom system, and, for sure, the original stimulus to study set theory was not to encompass a plethora of concepts of set membership. Secondly, it can be seen as an artificial attempt at reducing truth to provability by decree, while at the same time saying nothing about the justification of the axioms. Thirdly, in this single-layer view of set theory, we have no way to understand $\in$ as capturing a concept of membership: It is just an arbitrary relation which satisfies a given list of axioms of a formal system for set theory. In this case, not even the interpretation of finitary sentences is satisfactory, for different models, possibly non-wellfounded, may disagree with respect to these sentences. However, it is very reasonable to assume that we can interpret finitary sentences in a satisfactory way, for, otherwise, it would be very hard to see how we could understand the formal systems themselves. More dramatically, from Gödel's second incompleteness theorem it follows that the consistency of set theory itself is one of those finitary sentences that witnesses this shortcoming of interpretation in the single-layer account of set theory in question. Gödel himself remarked that:

> ... It is *this* theorem [the second incompleteness theorem] which makes the incompletability of mathematics particularly evident. For, *it makes it impossible that someone should set up a certain well-defined system of axioms and rules and consistently make the following assertion about it: All of these axioms and rules I perceive (with mathematical certitude) to be correct, and moreover I believe that they contain all of mathematics.* If somebody makes such a statement he contradicts himself. For if he perceives the axioms under consideration to be correct, he also perceives (with the same certainty) that they are consistent. Hence he has a mathematical insight not derivable from his axioms. ([9], p. 309)

Although I do not believe that mathematical knowledge can be extracted from perception, I agree with Gödel that the consistency of a mathematical theory is a presumed component of the thought that led to its formal systems. Therefore, I argue that a satisfactory account of the truth of the axioms of a formal system for a mathematical theory must also account for the truth of their consistency, if they are consistent, and, at the same time, leave some room for the possibility of inconsistency. In fact, our grasp of formal systems *presupposes* that we understand mathematical sentences of the form "S is provable from the axioms" either affirmed or denied. This is accomplished in the double-layer pictures of set theory that I am proposing. For, on the one hand, if the formal systems for set theory are consistent then the truth value of their consistency is fixed by both the first and the second lists of directive principles, as it was seen in Sect. 3. On the other hand, if the formal systems for set theory are inconsistent, then both lists of directive principles are *eo ipso* incoherent and fail to play any role as a criterion for the interpretation of the formal systems.

This possibility of failure is an important aspect of this account of set theory because, indeed, we may fail. How could we fail to formulate a consistent system of axioms and rules if we do have a nonpropositional grasp of a "Model-in-the-Sky", which is declared as the standard for mathematical correctness? How could our formal systems for set theory turn out to be inconsistent if we *perceive with mathematical certitude* the truth of their axioms? The directive principles do not give us certainty regarding consistency: They can only give us a criterion for interpreting the formal language such that if the consistency sentences are *assumed to be true*, then they are true according to the given interpretation, which is exactly what provability cannot accomplish according to Gödel's second incompleteness theorem. The delimitative role played by directive principles when demarcating the class of standard models of the theory and fixing truth-values is very different from the role of "mathematical intuition" as a prior standard that the axioms have to meet in order to be considered correct. In [7], Gödel argued for "mathematical intuition" as a criterion of truth in set theory different from provability in a formal setting:

> ... What, however, perhaps more than anything else, justifies the acceptance of this criterion of truth in set theory is the fact that continued appeals to mathematical intuition are necessary not only for obtaining unambiguous answers to the questions of transfinite set theory, but also for the solution of the problems of finitary number theory (of the type of Goldbach's conjecture), where the meaningfulness and unambiguity of the concepts entering into them can hardly be doubted. This follows from the fact that for every axiomatic system there are infinitely many undecidable propositions of this type. ([7], p. 485)

I do not think that an appeal to some obscure mathematical intuition clarifies anything. I think that a priori mathematical intuition is something we know nothing about. We know much more about mathematical truth than about a priori intuition: explaining the former in terms of the latter sounds like explaining the partially understood in terms of the completely not understood. If mathematical intuition is understood as a posteriori mathematical feeling, then it is something we *gain* through continued mathematical training. We could be trained in a mathematical subject until the basic articulation of its primitive notions become fully grasped and we forget how it was when we were not yet thinking that way. At this point, we could say that the fundamentals of the subject became intuitive to us, but to say that this a posteriori intuition is the foundation of the subject is to reverse the order of explanation. I am defending a replacement of 'mathematical intuition' by 'directive principles', which, of course, is not a change of names – their roles and nature are very different. With this proviso, I would agree with a thesis that is related to Gödel's extract above: As soon as we want to determine in what relation set theory stands to and to what extent it is captured by formal systems, then analysis of directive principles is essential. The reason behind this is that formal systems cannot provide the criteria for structures that directive principles can.

Trying to clarify what is meant by mathematical intuition, Gödel remarked earlier that:

> It should be noted that mathematical intuition need not be conceived of as a faculty giving an *immediate* knowledge of the objects concerned. Rather it seems that, as in the case of physical experience, we *form* our ideas also of those objects on the basis of something else which *is* immediately given. Only this something else here is *not*, or not primarily, the sensations. ([7], p. 484)

Gödel does not tell us what this "something else which is immediately given" is supposed to be, but there are some important remarks on the directive principles that are related to Gödel's attempt to clarify the issue of mathematical intuition: The directive principles are not a faculty giving an *immediate* knowledge of sets. Rather it seems that we form our ideas of the mathematical objects concerned on the basis of the directive principles in the sense that the subject matter of set theory is prescribed by these principles. They are an outcome of the thoughts of the founding fathers of this mathematical theory, but their origin is not a primary concern of the investigations on the possibility of interpreting formal systems for set theory. In any case, it is clear that they are not sensations.

Ferreirós ([5], pp. 380–384) is another logician-philosopher that has recently expressed a related view on the insufficiencies and implausibility of an account of mathematical theories according to which they consist of formal systems standing alone. When talking about two contrary tendencies in the development of mathematics, one which aims at a reduction of mathematics to a purely symbolic system and another one which aims at a reduction of mathematics to a purely conceptual system, Ferreirós writes

> In my view, the failure of both radical tendencies is of the essence. The standpoint I adopt emphasizes the need to consider the meaning or thought that accompanies formulae and calculations. (This is no doubt shared by many other philosophers, but the question is how to proceed.) Mathematical symbolism cannot be mastered without immersion in a practice, and by learning the practice we learn to associate representations and meaning to the formulae. Normal (so-called informal) symbolic systems and theories cannot be made to stand alone outside of practice; and when systems and theories are formalized and made to stand alone, the phenomenon of non-standard interpretations arises in a natural way.
>
> Indeed, I defend the *complementarity of symbolic means and thought* in mathematics–each one joined by the other, none of them reducible to the other. For obvious reasons, it is more difficult to deny the role and importance of the symbolic component in mathematics, but substantial arguments can be given for a similar conclusion concerning the conceptual component. For my purposes here, I shall be content with the modest claim that, in light of developments in mathematics and its foundations during the 20th century, such a standpoint deserves to be seriously considered as an option. ([5], pp. 381 and 382)

Of course, I agree with Ferreirós on the insufficiency of symbolic means to account for (all) mathematical theories. However, it is not enough to say that mathematics and set theory are a combination of a symbolic component with a conceptual component – if one is trying to formulate in these terms a standpoint on the foundations of mathematical truth then one must say what the symbolic and conceptual components are supposed to be and how they relate to each other. I am defending that, in the case of set theory, *directive principles*, which give unambiguous and objective criteria to demarcate the interpretation and understanding of formal systems, constitute the conceptual component[6] complementing the layer of formal systems, and not *thought*, *meaning* and *representations*, which are the categories used by Ferreirós in the above extract. Ferreirós does not tell us what exactly this conceptual component he is talking about is: It is not even clear whether this conceptual component is public or private. I do not think that the intentions, representations and thoughts that we associate to formulae when we learn a practice are univocal. The problem of the possibility of interpreting formal systems is in need for a more precise account of the conceptual component of the corresponding concrete axiom systems.

I am proposing the complementarity of formal systems and the directive principles – formal systems cannot play the delimitative role that the directive principles can, and these principles cannot play the deductive role of formal systems.

[6]The conception of the conceptual as a criterion is unproblematic: A concept naturally gives rise to a criterion separating those things falling under it from the rest.

3 Mathematical Analysis of the Role of Directive Principles

Although lists of directive principles do not constitute a formal system, their role as criteria for the interpretation of formal systems can be *analyzed* within mathematics. Before someone says this is circular, I reinforce that I am not talking about any *reduction* of directive principles to formal systems. Circularity comes in only if a reduction is proposed: It is circular to reduce directive principles to formal systems, since those principles are a basis for interpreting formal systems. However, it is *not* circular to use a formal system for set theory to *analyze* directive principles. There is no reduction taking place here, the layer of the formal systems does not subsume the layer of the directive principles and, consequently, there is no circularity.

From now on, I will assume that the formal system ZFC for set theory is consistent. The aim of a mathematical analysis of the role of directive principles is to clarify what is, according to each of the two lists of principles, the subject matter of set theory.

3.1 The Standard Models of Set Theory According to the First List of Directive Principles

I take it for granted that a mathematical analysis of the subject matter of an axiom system for a mathematical theory must result in a distinguished class of structures, the standard models. Therefore, the following definition is required:

Definition 1 A structure (M, E), in which E is a binary relation on M, is said to *conform* to the first list of directive principles iff

1. (M, E) satisfies the axioms of extensionality and regularity.
2. If $x \in M$ and if c is a subset of M such that for each $y \in c$ it holds that yEx, then there is an element in M whose E-members are the elements of c.
3. If $x \in M$ and if $r : M \to M$ is a function on M, then there is an element in M whose E-members are all those elements w such that $r(y) = w$ for some y that is E-member of x.
4. If $x \in M$ then there is an element in M whose E-members are all E-members of E-members of x.
5. If $x \in M$ then there is an element in M whose E-members are all E-subsets of x.
6. (M, E) satisfies the axiom of infinity.

Theorem 1 characterizes those structures conforming to the first list of directive principles, clarifying what is, according to those principles, the subject matter of set theory:

Theorem 1 (Zermelo [15]) *A structure* (M, E) *conforms to the first list of directive principles iff for some strongly inaccessible cardinal* κ, $(M, E) \cong (V_\kappa, \in)$.

Proof First, notice that E is a well-founded relation on M. In fact, suppose that there is an infinite sequence $(x_i)_{i\in\omega}$ of elements of M such that $x_{i+1}Ex_i$, for each $i \in \omega$. Since (M, E) satisfies all the axioms of ZFC, it follows that every element in M has a transitive closure in the sense of (M, E). Let x be the transitive closure of x_0 in (M, E), and let c be the set $\{x_i : i \in \omega\}$. From the second directive principle in the first list, it follows that there is a set $y \in M$ such that zEy iff there is an $i \in \omega$ such that $z = x_i$. Thus,

$$(M, E) \models \forall z(zEy \rightarrow (\exists w(wEy \wedge wEz))),$$

and (M, E) cannot satisfy the axiom of regularity. Therefore, it is not the case that there is an infinite sequence $(x_i)_{i\in\omega}$ of elements of M such that $x_{i+1}Ex_i$, for each $i \in \omega$.

Now, let $(N, \in)$ be the transitive collapse of (M, E). If α is the first ordinal which is not in N, then $N \subseteq V_\alpha$. Suppose $N \neq V_\alpha$. The ordinal α is a limit ordinal. For if α is $\beta + 1$, then $\beta \in N$ and the set $\beta \cup \{\beta\} \in V_{\alpha+1} \setminus V_\alpha$ cannot be in N. From this it follows that $(N, \in)$ cannot be a transitive model of ZFC and we conclude that α is limit.

Let β be the first ordinal such that there is an element x in $V_\beta \setminus N$. Since α is limit, $V_\alpha = \bigcup_{\gamma<\alpha} V_\gamma$, and if for all $\gamma < \alpha$ it holds that $V_\gamma \subseteq N$, then $V_\alpha \subseteq N$. Therefore $\beta < \alpha$ and $\beta \in N$. Since $x \subseteq N \cap V_\beta = (V_\beta)^N \in N$, it follows, from the second directive principle, that $x \in N$. Therefore, $N = V_\alpha$.

From the third and fifth directive principles, it can be proved that α is regular and strong limit, respectively. In fact, if β is an ordinal such that $\beta < \alpha$, then $\beta \in N$ and the direct image of a function $r : \beta \rightarrow N$ is an element of N. Since the union of an element in N is also in N, it follows that $r : \beta \rightarrow N$ cannot be cofinal in α. Also, if $\lambda < \alpha$ is a cardinal, then $\wp(\lambda) \in N$, the cardinal of $\wp(\lambda)$ in the sense of $(N, \in)$ is 2^λ, and $2^\lambda < \alpha$. This proves the result. □

From this mathematical result, it is plausible to say that, according to the first list of directive principles, set theory is about the concept of the cumulative hierarchy of sets, which is exemplified by the hierarchies of sets $(V_\kappa, \in)$, in which κ is a strongly inaccessible cardinal. Now, is this the correct interpretation of set theory? Are the hierarchies of sets the *true* subject matter of set theory? I think that the only sure answer to this question is that this interpretation is historically correct, in the sense that (*a*) the concept of the cumulative hierarchy of sets seems to be the basic concept in the discussions on set-theoretic truth, and (*b*) the directive principles in the first list are to be found as the basic facts about sets in the works of the founding fathers of this mathematical theory, and fixes truth-values in a plausible way, as it will be shown in the sequel. Therefore, although directive principles are not intentions, they seem to fulfill the intentions of the framers of the axiom systems for set theory, notably Cantor's and Zermelo's intentions.

Now it is possible to clarify how sentences in formal systems for set theory acquire a meaning under the background language convention given by the first list of directive principles. Adopting a model-theoretic perspective on this point, if φ is a sentence (in the first-order language with a symbol for equality and one binary predicate variable R for membership) then the *meaning* of φ can be defined, according to the background language convention given by the first list of directive principles, as the class function that assigns to each structure $(V_\kappa, \in)$, in which κ is a strongly inaccessible cardinal, the truth-value **T** if $(V_\kappa, \in) \models \varphi$, and the truth-value **F** otherwise. This extensional account of meaning is based on the thesis that a sentence expresses its truth-conditions with respect to a presumed background language convention. In this model-theoretic setting, the truth-conditions of a sentence φ in a formal system for set theory can be identified with structures $(V_\kappa, \in)$ in which φ is true. The truth-value of φ is said to be *fixed by the directive principles* if and only if φ has the same truth-value on every structure $(V_\kappa, \in)$, in which κ is a strongly inaccessible cardinal. It is already easy to see that the directive principles and the first-order axioms of the formal system ZFC have very different fixing powers.[7] If there is an explanation to this greater fixing powers of the directive principles, then it is that, contrary to the axioms of a formal system, the directive principles are *not* required to effectively generate the truths that are fixed by them. The role of the directive principles is not to computably generate truths, but just to give a criterion for the interpretation of formal systems.

Recall that, in a model-theoretic perspective, if φ is a sentence then the truth-value of φ is fixed by the background language convention given by the first list of directive principles iff φ has the same truth-value on every structure $(V_\kappa, \in)$, in which κ is a strongly inaccessible cardinal, that is, iff the class function which is the meaning of φ is constant. The structures (isomorphic to) $(V_\kappa, \in)$, in which κ is a strongly inaccessible cardinal will be called *Z-standard models* of ZFC. The continuum hypothesis, CH, for example, has the same truth-value on every Z-standard model[8] because all sets relevant to the truth or falsity of CH belong to a lower level V_α, in which α is countable. Therefore, the truth-value of CH is fixed by the directive principles. Also, the truth-value of arithmetical sentences – consistency statements, in particular – are fixed[9] by the first list of directive principles, and, of course, the truth of every theorem of ZFC is fixed by the first list of directive principles.

The mathematical analysis of the truth-value of some statements can be conditional. Consider the sentence ‘there is a strongly inaccessible cardinal’. If there is a strongly inaccessible cardinal, then this sentence does not have the same truth-value on every Z-standard model. If there are no strongly inaccessible cardinals then the falsity of the sentence ‘there is a strongly inaccessible cardinal’ is fixed by the first

[7] Naturally, the truth-value of φ is said to be fixed by the axioms of ZFC iff φ has the same truth-value on every model of ZFC.

[8] Notice that this does not mean that CH holds in every Z-standard model. It just means that it is true/false in one Z standard model iff it is true/false in all Z-standard models.

[9] From Gödel’s first incompleteness theorem it follows that the sentences that hold in all standard models cannot be effectively enumerated. This shows that the arbitrariness present in directive principles (2) and (3) of the first list cannot be fully formalized.

list of directive principles. Although this is not as informative as for CH, it gives some information: If 'there is a strongly inaccessible cardinal' is true then its truth is not fixed by the first list of directive principles, and if it is false then its falsity is fixed by the first list of directive principles. Similarly, the existence of a measurable cardinal can be analyzed conditionally: If 'there is a measurable cardinal' is true then its truth is not fixed by the first list of directive principles, and if it is false then its falsity is fixed by the first list of directive principles.

Another example of a conditional analysis is the following: If $V = L$ holds then it holds in every Z-standard model and its truth is fixed by the directive principles. On the other hand, if $V = L$ is false, then its falsity can be fixed or not by the directive principles depending on how it fails: If there is a non-constructible subset of ω then $V = L$ is false in every structure $(V_\kappa, \in)$ and its falsity is fixed by the first list of directive principles. If there is a strongly inaccessible cardinal and the rank of the least-ranked non-constructible sets is greater than the first strongly inaccessible cardinal, then the falsity of $V = L$ is not fixed by the first list of directive principles. Therefore, it seems that the first list of directive principles fix truth-values in a satisfactory way, meeting *desiderata* (iv). For example, the directive principles fix the truth-value of every arithmetic statement, and every statement whose truth or falsity depends only on an initial segment of the hierarchies of sets bounded by the first inaccessible. It is also plausible to say that the existence of an inaccessible cardinal, and related statements, does not have the truth-value fixed.

If the truth of a sentence assumed to be true is not fixed by the directive principles, then its adoption as a new axiom for set theory must be accompanied by an extension of directive principles, in case one wants to keep the possibility of interpreting the resulting formal systems. If the truth of a new axiom assumed to be true is not in the range of the directive principles stated above, then these principles are insufficient for interpreting the new axiom system. On the other hand, if the truth of a sentence assumed to be true is fixed by the directive principles then its adoption as a new axiom need not be accompanied by an extension of directive principles, but this does not mean that its adoption as a new axiom is *justified* in the picture of set theory that I am defending. Because, even in case the truth of a sentence assumed to be true is fixed by the directive principles, its adoption as a new axiom may cause a mismatch between the layer of the directive principles and the layer of formal systems. In fact, if $V = L$ is true then its truth is fixed by the directive principles, but $V = L$ cannot be obtained by a formalization of part of the directive principles, which implies that $ZF + V = L$ is not a formal system for set theory according to the double-layer picture of set theory presented here. Therefore, in this double-layer scheme, the mere conditional fact that if a sentence is true then its truth is fixed by the directive principles does *not* suffice to justify the adoption of the sentence as a new axiom. In this schema, it is also required that the resulting formal systems be obtained from formalizations of the directive principles.

3.2 The Standard Models of Set Theory According to the Second List of Directive Principles

The mathematical analysis of the criterion given by the second list of directive principles is based on the following:

Definition 2 A structure (M, E), in which E is a binary relation on M, is said to *conform* to the second list of directive principles iff

1. (M, E) satisfies the axioms of ZFC.
2. If c is a countable subset of M, then there is an element in M whose E-members are the elements of c.

It is easy to prove that a structure conforms to the second list of directive principles iff it is isomorphic to a transitive model of ZFC closed under countable subsets.

Theorem 2 *A structure (M, E) conforms to the second list of directive principles iff (M, E) is isomorphic to a transitive model of ZFC closed under countable subsets.*

Proof If (M, E) conforms to the second list of directive principles, then it is well-founded. In fact, suppose that there is an infinite sequence $(x_i)_{i\in\omega}$ of elements of M such that $x_{i+1} E x_i$, for each $i \in \omega$. From the second item in Definition 2, it follows that there is a set $y \in M$ such that zEy iff there is an $i \in \omega$ such that $z = x_i$. Therefore, (M, E) cannot satisfy the regularity axiom.

Now, let $(N, \in)$ be the transitive collapse of (M, E). The transitive structure $(N, \in)$ is isomorphic to (M, E), and it conforms to the second list of directive principles. It is a model of ZFC and it is closed under countable subsets.

The other direction is trivial. □

Since all transitive models of ZFC agree on arithmetic statements, the second list of directive principles fixes the truth-values for all arithmetical sentences, which is an important desideratum for a criterion for the interpretation of ZFC. The status of the continuum hypothesis is less determined. The power set of ω is absolute for transitive models which are closed under countable subsets. Also, $\aleph_1$ is absolute for those models. Therefore, if $\aleph_1$ is the cardinal of $\wp(\omega)$ in one transitive model closed under countable sets, then it is really the cardinal of $\wp(\omega)$. Equivalently, if the continuum hypothesis is false, then its falsehood is fixed by the second list of directive principles.

The notion of countable set is also absolute for transitive models of ZFC closed under countable subsets, and it seems sensible to say that the absoluteness of the notion of countability is part of the very conception of set theory.

In the proposed double-layer schema, the conceptual component of set theory is given in the first layer as a criterion separating some structures from the class of all models of the corresponding formal systems. Theorem 2 above shows that according to the second list of directive principles the subject matter of set theory can be understood as constituted by the transitive models closed under countable subsets, which can be called K-standard models. This is probably not the usual way the conceptual component of set theory is understood, but this account of set-theoretic truth has the merit of fixing the truth-values of arithmetic statements, and even those of second-order arithmetic, because of the absoluteness of $\wp(\omega)$ with respect to K-standard models, and its semantic commitments are relatively modest.

3.3 The Standard Models of Elementary Arithmetic

It is generally acknowledged that (most) mathematical theories can be faithfully reduced to a definitional extension of set theory. If this is accepted, then the analysis of set theory developed in Sects. 2 and 3 is perfectly general. However, this is not an unproblematic thesis. For example, there are the usual, well-known arguments to the effect that set theory cannot provide an ontological reduction of mathematics.[10] Therefore, an independent account of mathematical truth for other theories is desirable. Fortunately, it is possible to provide a similar double-layer account to other foundational axiomatic theories, such as elementary arithmetic,[11] an axiomatic theory in which truth and provability also seem to be mismatched. In contrast with the theory of groups, for example, in which permutation groups are the primary phenomena, the (first-order) formal systems for set theory and elementary arithmetic cannot be obtained as axiomatizations of such classes of models that we come upon as the primary phenomena, independently of those formal systems. Instead, the formal systems for set theory and elementary arithmetic can be obtained as formal counterparts of directive principles. In the case of elementary arithmetic the double-layer schema also applies — the first layer consists of the directive principles, the second layer consists of formal systems. A list of directive principles for elementary arithmetic, formulating the notational-algorithmic conception of arithmetical operations, can be given as follows:

1. Each number is denoted by a unique numeral, which is a syntactic object obtained by a repetition, possibly null, of a primitive symbol. Each numeral denotes a unique number.
2. Given two nsumerals s and t, the sum of the numbers denoted by s and t is denoted by the numeral obtained by the repetition of the primitive symbol determined by t over s.
3. Given two numerals s and t, the product of the numbers denoted by s and t is denoted by the numeral obtained by the repetition of the repetition s determined by t.

[10]For a clear and concise exposition of this point, see [10], pp. 36–39.

[11]I am using the expression “elementary arithmetic” to designate the axiomatic theory of natural numbers without reference to sets of natural numbers. In order to designate the axiomatic theory of natural numbers and sets of natural numbers the expression “elementary analysis” is preferred.

The subject matter of elementary arithmetic can be analyzed in an analogous way: It will be proved that all structures conforming to the principles above in the obvious sense are isomorphic to the structure $(\omega, +, \times)$. Therefore, that elementary arithmetic is concerned with the standard model is, in this account, a consequence of the very directives which are the conceptual component of this axiom system, and this is very different from declaring a model standard by decree.

These directive principles prescribe a criterion for the interpretation of formal systems for elementary arithmetic. Now, there is a mathematical theorem which can be seen as evidence for the thesis that in this specific case the directive principles do what is expected, and their first-order formalizations are always *partial*:

Theorem 3 *Every structure $\mathscr{A} = (D, \oplus, \otimes)$ conforming to the directive principles above, in the sense that:*

1. *for every $d \in D$ there is a unique numeral s such that $d = s^{\mathscr{A}}$,*
2. *given two elements $s^{\mathscr{A}}$ and $t^{\mathscr{A}}$ in D, the sum $s^{\mathscr{A}} \oplus t^{\mathscr{A}} = (s\hat{\ }t)^{\mathscr{A}}$, where $s\hat{\ }t$ is the numeral obtained by the repetition of the primitive symbol determined by t over s, and*
3. *given two elements $s^{\mathscr{A}}$ and $t^{\mathscr{A}}$ in D, the product $s^{\mathscr{A}} \otimes t^{\mathscr{A}} = (s * t)^{\mathscr{A}}$, where $s * t$ is the numeral obtained by the repetition of the repetition s determined by t, is isomorphic to the structure $(\omega, +, \times)$.*

Proof For each $d \in D$, let s be the unique numeral such that $d = s^{\mathscr{A}}$, and let n be the unique number canonically associated with s. Clauses (2) and (3) imply that the bijection $d \mapsto n$ is an isomorphism. □

The directive principles given for elementary arithmetic can be equivalently expressed in $L_{\omega_1\omega}$. Therefore its semantic commitments are very modest since this language is arguably a slight extension of the first-order language of elementary arithmetic.

4 On Foundations of Mathematical Truth

It is usually said that set theory is a *foundation of mathematics*. I understand this as the claim: Mathematical truth can be explained away in terms of the better understood set-theoretic truth. Thus, according to this understanding of the above claim, set theory is a foundation of mathematical truth. I agree with this claim, but this is not an unproblematic thesis. In fact, as it was already mentioned, there are well-known arguments to the effect that set theory cannot provide an ontological reduction of mathematics. However, an answer to this objection is that to provide an ontological reduction is not a requirement for a foundation of mathematical truth: What is required for set theory to be a foundation of mathematical truth is that (*a*) all mathematical theorems can be formalized as theorems in the formal systems for set theory, and (*b*) this mathematical theory comes with a good understanding of its truth. Item (*a*) hardly needs to be argued for. With respect to (*b*), if the explanation of mathematical truth in terms of set-theoretic truth is to represent an improvement in clarity and precision, then set-theoretic truth must be well-understood.

In this paper I have defended an understanding of the commitments behind set-theoretic truth based on certain directive principles. These directives were subject to a mathematical analysis in order to clarify their delimitative role in the interpretation of set theory. In that mathematical analysis, the subject matter of set theory is, according to the first list of directive principles and Theorem 1, constituted by the hierarchies of sets $(V_\kappa, \in)$, in which κ is a strongly inaccessible cardinal, and according to the second list and Theorem 2, is constituted by the transitive models of ZFC which are closed under countable subsets. These were established as mathematical facts *based on Definitions* 1 and 2, respectively. Since set theory is considered a standard foundation of mathematics, these accounts of the commitments behind the interpretation of set theory correspond to accounts of the commitments behind mathematical truth. Now, it is important to finally explain how these pictures of set theory and mathematical truth solve the original problems.

The presumed concrete character of formal systems for set theory must be defined. Based on Gödel's incompleteness theorems, set-theoretic truth is generally held to transcend formal systems. If set-theoretic truth is supposed to go beyond provability in formal systems, then we have to explain how can we ground truth-values in set theory and what are we committed to when we do this. Our first task is to define what is a concrete axiom system. The proposed solution to this problem was formulated in terms of a double-layer schema of concrete axiom systems: A concrete axiom system is constituted by two layers, the layer of directive principles of the system, which correspond to the conceptual component of the system, and the layer of the formal systems, which correspond to the deductive component. Directive principles give us the commitments latent in the interpretation of formal systems, separating appropriate structures from all possible realizations of the axioms (when mathematically analyzed).

The second thing to do is to give directive principles for set theory. Two lists of directive principles were presented. The double-layer pictures of set theory and set-theoretic truth unfolded in Sects. 2 and 3 are such that truth-values are fixed beyond provability. For example, each sentence of second-order arithmetic has a rigid truth-value in each one of the double-layer pictures presented here. Of course, the truth of every theorem of 20th century classical mathematics which can be formalized in ZFC is also fixed in these double-layer pictures. That is, transferring the analysis of the commitments behind set-theoretic truth set forth in this paper to 20th century classical mathematics, it is possible to provide an appropriate account of the commitments latent in the notion of mathematical truth we happen to have inherited that does not rest on mathematical intuition, nor on formalization alone, but on principles which give us criteria for interpreting formal systems.

The proposed double-layer schema of concrete axioms systems is not committed to objects in an independent model, but it is committed to the objectivity of the criterion given by directive principles in the first layer, which corresponds to the conceptual component of the system. A given structure either conforms to the criterion or not, according to the relevant definitions of conformity, and this is what is understood by the objectivity of the criterion given by directive principles. It is a semantic commitment, it is what is required by the conceptual component of the system in order to separate some structures from all possible realizations of the axioms. The second list of directive principles given for set theory has a modest semantic commitment, when compared to the first list, but the first list seems to be closer to the concept of membership hierarchy, which is arguably the most popular candidate for the conceptual component of set theory. I agree with Kreisel's dictum[12] that the point is not the existence of objects but the objectivity of mathematical truth. Philosophy of mathematics can profit from a shift of focus from the category of the object to the notion of objectivity, and to account for set-theoretic truth in terms of an objective criterion is such a shift.

References

1. Bernays, P. 1942. *A System of Axiomatic Set Theory: Part III. Infinity and Enumerability. Analysis*. The Journal Of Symbolic Logic.
2. Cantor, G. 1899. *Letter to Dedekind*, [translated in van Heijenoort 1967, 113–117].
3. Cantor, G. 1955. *Contributions to the Founding of the Theory of Transfinite Numbers*. Dover.
4. Cohen, P. 2008. *Set Theory and the Continuum Hypothesis*. Dover.
5. Ferreirós, J. 2012. *On Arbitrary Sets and ZFC*, Bulletin of Symbolic Logic.
6. Fraenkel, A., Y. Bar-Hillel, and A. Levy. 1973. *Foundations of Set Theory*. North-Holland.
7. Gödel, K. 1947, 1963. *What is Cantor's Continuum Problem* [in Benacerraf and Putnam 1983, 470–485].
8. Gaifman, H. 2012. *On Ontology and Realism in Mathematics*. Review of Symbolic Logic.
9. Gödel, K. 1995. *Collected Works*, vol. III. Oxford.
10. McGee, V. 1997. *How Can we Learn Mathematical Language*, The Philosophical Review, vol. 106, no. 1, pp. 35–68.
11. Moore, G. 2013. *Zermelo's Axiom of Choice*. Dover.
12. Shoenfield, J. 2001. *Mathematical Logic*. ASL.
13. Tait, W. 2005. *The Provenance of Pure Reason*. Oxford.
14. Zermelo, E. 1908. *Untersuschungen über due Grundlagen der Mengenlehre, I*, Mathematische Annalen 65, 261–281 [translated in van Heijenoort 1967, 199–215].
15. Zermelo, E. 1930. Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae* 16: 29–47.

[12]Folklore. See [8], footnote 4.