

# What Mathematical Theories of Truth Should be Like (and Can be)

Seppo Heikkilä

Department of Mathematical Sciences, University of Oulu

BOX 3000, FIN-90014, Oulu, Finland

E-mail: sheikki@cc.oulu.fi

November 3, 2014

## Abstract

Hannes Leitgeb formulated eight norms for theories of truth in his paper: ‘What Theories of Truth Should be Like (but Cannot be)’. We shall present in this paper a mathematical theory of truth for suitably constructed languages which contain the first-order language of set theory, and prove that it satisfies all those norms.

**MSC:** 00A30, 03B10, 03C62, 03D60, 03F50, 06A07, 47H04, 47H10

**Keywords:** Set theory, Zermelo-Fraenkel, model, first-order, theory, language, sub-language, formal, sentence, truth, fixed point, chain, transfinite sequence.

## 1 Introduction

Hannes Leitgeb formulated in [19] the following norms for theories of truth:

- (a) Truth should be expressed by a predicate (and a theory of syntax should be available).
- (b) If a theory of truth is added to mathematical or empirical theories, it should be possible to prove the latter true.
- (c) The truth predicate should not be subject to any type restrictions.
- (d)  $T$ -biconditionals should be derivable unrestrictedly.
- (e) Truth should be compositional.

- (f) The theory should allow for standard interpretations.
- (g) The outer logic and the inner logic should coincide.
- (h) The outer logic should be classical.

We present a mathematical theory of truth (shortly MTT) that satisfies norms (a)–(h). It is formulated for suitably constructed languages which contain the first-order language  $L = \{\in\}$  of set theory. They are sublanguages of  $\mathcal{L} = \{\in, T\}$ , where  $T$  is a monadic predicate. While not entailing all formalities of first-order languages, they are closed under logical connectives and quantifiers of first-order predicate logic.

Zermelo-Fraenkel (ZF) set theory, sets of its minimal model  $M$ , concepts definable in  $L$  and classical two-valued logic are main tools. MTT proves ZF set theory modelled by  $M$  true. As stated in [14], “First-order Zermelo-Fraenkel set theory is widely regarded as the standard of mathematical correctness, in the sense that a proof is correct if and only if it can be formalised as a formal proof in set theory.”

As shown in [1], “with some elementary assumptions about truth and logic, a logical disaster ensues.” It is caused by Liar paradox. Paradoxes led Zermelo to axiomatize set theory. By [19] Tarski “excluded all Liar-like sentences from being well-formed” from his theory of truth (cf. [22]). Construction of sublanguages enables MTT to satisfy also the following norms:

- (i) The mathematical theory of truth should be free from paradoxes.
- (j) Truth should be defined for languages within the very languages themselves.

MTT is immune to ‘Tarski’s Commandment’ (cf. [20]), to Tarski’s Undefinability Theorem (cf. [22]), to Tarskian hierarchies (cf. [8]), and to ‘Revenge of the Liar’ (cf. [5]).

Model  $M$  can be replaced by any countable model of ZF or other set theory. Examples of such theories are presented, e.g., in [18]. The language  $L$  can be replaced by the language of arithmetic, and  $M$  by a standard model of arithmetic. In [13] MTT is extended to a more general class of languages.

PKF theory (cf. [9, 10]) axiomatizes Kripke’s theory of truth (cf. [17]). Both inner and outer logic of PKF are partial, whence norms (h) and (j) are not valid. Kripke’s theory don’t satisfy norms (d) and (g) because of partially interpreted truth predicate and three-valued inner logic.

Theory KF, which axiomatizes Kripke’s closed-off models (cf. [4, 5, 17]), don’t satisfy norms (d), (g) and (i). Norm (g) is rejected in [5]. Preferring norm (g) over any mixture of two different logics, and (h) over partial logic, we use classical logic and construct sublanguages. Modifying Wittgenstein: Whereof one must not speak, thereof one can be silent.

## 2 Preliminaries

As for the used terminology, cf. e.g., [12, 18]. Denote by  $M$  the minimal, countable, transitive and well-founded model of ZF set theory constructed in [3]. It follows from [18, Lemma II.8.22] that for each sentence  $A$  of the language  $L = \{\in\}$  either  $A$  or its negation  $\neg A$  is true in  $M$  in the sense defined in [18] (cf. [18, II.7 and p. 237]).  $M$  contains by [3, 16] natural numbers. Assume that their names, called numerals, are defined in  $L$ , e.g., as in [7].

First we shall construct a family of sublanguages for the language  $\mathcal{L} = \{\in, T\}$ , where  $T$  is a monadic predicate. Denote by  $D$  the set of Gödel numbers of sentences of  $\mathcal{L}$  in its fixed Gödel numbering. Let  $\#A$  denote the Gödel number of a sentence  $A$  of  $\mathcal{L}$ , and  $\lceil A \rceil$  the numeral of  $\#A$ . Given a subset  $U$  of  $D$ , define subsets  $G(U)$  and  $F(U)$  of  $D$  by following rules (iff abbreviates if and only if):

- (r1) If  $A$  is in  $L$ , then  $\#A$  is in  $G(U)$  iff  $A$  is true in  $M$ , and in  $F(U)$  iff  $\neg A$  is true in  $M$ .
- (r2) If  $A$  is in  $\mathcal{L}$ , then  $\#T(\lceil A \rceil)$  is in  $G(U)$  iff  $\#A$  is in  $U$ , and in  $F(U)$  iff  $\#[\neg A]$  is in  $U$ .

Sentences determined by rules (r1) and (r2), i.e., all sentences  $A$  of  $L$  and those sentences  $T(\lceil A \rceil)$  of  $\mathcal{L}$  for which  $\#A$  or  $\#[\neg A]$  is in  $U$ , are called *basic sentences*.

Next five rules deal with logical connectives. Let  $A$  and  $B$  be sentences of  $\mathcal{L}$ .

- (r3) Negation rule:  $\#[\neg A]$  is in  $G(U)$  iff  $\#A$  is in  $F(U)$ , and in  $F(U)$  iff  $\#A$  is in  $G(U)$ .
- (r4) Disjunction rule:  $\#[A \vee B]$  is in  $G(U)$  iff  $\#A$  or  $\#B$  is in  $G(U)$ , and in  $F(U)$  iff  $\#A$  and  $\#B$  are in  $F(U)$ .
- (r5) Conjunction rule:  $\#[A \wedge B]$  is in  $G(U)$  iff  $\#[\neg A \vee \neg B]$  is in  $F(U)$  iff (by (r3) and (r4)) both  $\#A$  and  $\#B$  are in  $G(U)$ . Similarly,  $\#[A \wedge B]$  is in  $F(U)$  iff  $\#[\neg A \vee \neg B]$  is in  $G(U)$  iff  $\#A$  or  $\#B$  is in  $F(U)$ .
- (r6) Implication rule:  $\#[A \rightarrow B]$  is in  $G(U)$  iff  $\#[\neg A \vee B]$  is in  $G(U)$  iff (by (r3) and (r4))  $\#A$  is in  $F(U)$  or  $\#B$  is in  $G(U)$ .  $\#[A \rightarrow B]$  is in  $F(U)$  iff  $\#[\neg A \vee B]$  is in  $F(U)$  iff  $\#A$  is in  $G(U)$  and  $\#B$  is in  $F(U)$ .
- (r7) Biconditionality rule:  $\#[A \leftrightarrow B]$  is in  $G(U)$  iff  $\#A$  and  $\#B$  are both in  $G(U)$  or both in  $F(U)$ , and in  $F(U)$  iff  $\#A$  is in  $G(U)$  and  $\#B$  is in  $F(U)$ , or  $\#A$  is in  $F(U)$  and  $\#B$  is in  $G(U)$ .

When  $A(x)$  is a formula in  $L$ , then quantifications  $\exists xA(x)$  and  $\forall xA(x)$  are sentences of  $L$ . Hence rule (r1) is applicable for them. So it suffices to set rules for  $\exists xT(x)$  and  $\forall xT(x)$ . Assume that the set  $X$  of numerals of Gödel numbers of sentences of  $\mathcal{L}$  is the intended domain of discourse for  $T$ . We set the following rules:

(r8)  $\#[\exists xT(x)]$  is in  $G(U)$  iff  $\#T(\mathbf{n})$  is in  $G(U)$  for some  $\mathbf{n} \in X$ , and  $\#[\exists xT(x)]$  is in  $F(U)$  iff  $\#T(\mathbf{n})$  is in  $F(U)$  for every  $\mathbf{n} \in X$ .

(r9)  $\#[\forall xT(x)]$  is in  $G(U)$  iff  $\#T(\mathbf{n})$  is in  $G(U)$  for every  $\mathbf{n} \in X$ , and  $\#[\forall xT(x)]$  is in  $F(U)$  iff  $\#T(\mathbf{n})$  is in  $F(U)$  at least for one  $\mathbf{n} \in X$ .

Rules (r1)–(r9) and induction on the complexity of formulas determine uniquely subsets  $G(U)$  and  $F(U)$  of  $D$  whenever  $U$  is a subset of  $D$ . Sublanguages  $\mathcal{L}_U$  of  $\mathcal{L}$  formed by those sentences  $A$  of  $\mathcal{L}$  for which  $\#A$  is in  $G(U)$  or in  $F(U)$  contain all sentences of  $L$  and are closed under logical connectives and quantifiers.

We say that a subset  $U$  of  $D$  is *consistent* iff both  $\#A$  and  $\#[\neg A]$  are not in  $U$  for any sentence  $A$  of  $\mathcal{L}$ . For instance, the empty set  $\emptyset$  is consistent.

The following two lemmas have counterparts in [11].

**Lemma 2.1.** *Let  $U$  be a consistent subset of  $D$ . Then  $G(U) \cap F(U) = \emptyset$ .*

*Proof.* If  $A$  is in  $L$ , then by rule (r1) and [18, Lemma II.8.22],  $\#A$  is not in  $G(U) \cap F(U)$ . By rule (r2)  $\#T(\lceil A \rceil)$  is in  $G(U)$  iff  $\#A$  is in  $U$ , and in  $F(U)$  iff  $\#[\neg A]$  is in  $U$ . Thus  $\#T(\lceil A \rceil)$  cannot be both in  $G(U)$  and in  $F(U)$  because  $U$  is consistent.

Make an induction hypothesis:

(h0)  $A$  and  $B$  are such sentences of  $\mathcal{L}$  that neither  $\#A$  nor  $\#B$  is in  $G(U) \cap F(U)$ .

As shown above, (h0) holds if  $A$  and  $B$  are basic sentences.

If  $\#[\neg A]$  is in  $G(U) \cap F(U)$ , then  $\#A$  is in  $F(U) \cap G(U)$ . Hence, if (h0) holds, then  $\#[\neg A]$  is not in  $G(U) \cap F(U)$ .

If  $\#[A \vee B]$  is in  $G(U) \cap F(U)$ , then  $\#A$  or  $\#B$  is in  $G(U)$ , and both  $\#A$  and  $\#B$  are in  $F(U)$  by (r4), so that  $\#A$  or  $\#B$  is in  $G(U) \cap F(U)$ . Hence, if (h0) holds, then  $\#[A \vee B]$  is not in  $G(U) \cap F(U)$ .

$\#[A \wedge B]$  cannot be in  $G(U) \cap F(U)$ , for otherwise both  $\#A$  and  $\#B$  are in  $G(U)$ , and at least one of  $\#A$  and  $\#B$  is in  $F(U)$ , contradicting with (h0).

If  $\#[\neg A]$  is in  $G(U) \cap F(U)$ , then  $\#A$  is in  $F(U) \cap G(U)$ , and (h0) is not valid. Thus, under the hypothesis (h0) neither  $\#[\neg A]$  nor  $\#B$  is in  $G(U) \cap F(U)$ . This result and the above result for disjunction imply that  $\#[\neg A \vee B]$ , or equivalently,  $\#[A \rightarrow B]$ , is not in  $G(U) \cap F(U)$ . Similarly,  $\#[A \leftrightarrow B]$  is not in  $G(U) \cap F(U)$ , for otherwise,  $\#A$  or  $\#B$  would be in  $G(U) \cap F(U)$  by rule (r7), contradicting with (h0).

It remains to show that  $\#[\exists xT(x)]$  and  $\#[\forall xT(x)]$  are not in  $G(U) \cap F(U)$ .

If  $U$  is empty, then  $T(\mathbf{n})$  is by rule (r2) neither in  $G(U)$  nor in  $F(U)$  for any  $\mathbf{n} \in X$ . Thus  $\#[\exists xT(x)]$  is by rule (r8) neither in  $G(U)$  nor in  $F(U)$ , and hence not in  $G(U) \cap F(U)$ .

If  $U$  is not empty, then  $\#A$  is in  $U$  for some  $A$  in  $\mathcal{L}$ . Since  $U$  is consistent, then  $\#[\neg A]$  is not in  $U$ . By rule (r2),  $\#T(\lceil A \rceil)$  is not in  $F(U)$ . Thus  $\#[\exists xT(x)]$  is by rule (r8) not in  $F(U)$ , and hence not in  $G(U) \cap F(U)$ .

Because  $U$  is consistent, it is a proper subset of  $D$ . Thus there is  $n \in D$  such that  $n \notin U$ . But  $n = \#A$  for some sentence  $A$  of  $\mathcal{L}$ , whence  $\#T(\mathbf{n}) = \#T(\lceil A \rceil)$  is not in  $G(U)$  by rule (r2). Consequently,  $\#[\forall xT(x)]$  is by rule (r9) not in  $G(U)$ , and hence not in  $G(U) \cap F(U)$ . The above results and induction on the complexity of formulas imply that  $\#A$  is not in  $G(U) \cap F(U)$  for any sentence  $A$  of  $\mathcal{L}$ .  $\square$

**Lemma 2.2.** *If  $U$  is a consistent subset of  $D$ , then both  $G(U)$  and  $F(U)$  are consistent.*

*Proof.* If  $G(U)$  is not consistent, then there is such a sentence  $A$  of  $\mathcal{L}$ , that  $\#A$  and  $\#[\neg A]$  are in  $G(U)$ . Because  $\#[\neg A]$  is in  $G(U)$ , then  $\#A$  is also in  $F(U)$  by rule (r3), and hence in  $G(U) \cap F(U)$ . But then, by Lemma 2.1,  $U$  is not consistent. Consequently, if  $U$  is consistent, then  $G(U)$  is consistent. The proof that  $F(U)$  is consistent if  $U$  is, is similar.  $\square$

### 3 A mathematical theory of truth and its properties

Recall that  $D$  denotes the set of Gödel numbers of sentences of the language  $\mathcal{L} = \{\in, T\}$ . Given a subset  $U$  of  $D$ , let  $G(U)$  and  $F(U)$  be the subsets of  $D$  constructed in Section 2.

**Definition 3.1.** *Assume that  $U$  is a consistent subset of  $D$ , and that  $U = G(U)$ . Denote by  $\mathcal{L}_U$  the language containing those sentences  $A$  of  $\mathcal{L} = \{\in, T\}$  for which  $\#A$  is in  $G(U)$  or in  $F(U)$ . A sentence  $A$  of  $\mathcal{L}_U$  is interpreted as true iff  $\#A$  is in  $G(U)$ , and as false iff  $\#A$  is in  $F(U)$ .  $T$  is called a truth predicate for  $\mathcal{L}_U$ .*

The existence of consistent subsets  $U$  of  $D$  satisfying  $U = G(U)$ , including the smallest one, is proved in Section 4.

The main goal of this paper is to prove the following Theorem for the above defined mathematical theory of truth (shortly MTT).

**Theorem 3.1.** *Assume that  $U$  is a consistent subset of  $D$ , and that  $U = G(U)$ . Then MTT satisfies all the norms (a)–(j) presented in the Introduction.*

*Proof.* (a) The syntax of MTT is comprised by logical symbols of first-order predicate logic, nonlogical symbols  $\in$  and  $T$ , natural numbers as constants, numerals as terms, and variables ranging in  $M$ .  $T$  is the truth predicate for  $\mathcal{L}_U$ , by Definition 3.1.

(b) Assume that  $A$  is a sentence of  $L$ . By [18, Lemma II.8.22] either  $A$  is true in  $M$ , or  $\neg A$  is true in  $M$  ( $A$  is false in  $M$ ).

- $A$  is true in  $M$  iff  $\#A$  is in  $G(U)$ , by rule (r1), iff  $\#A$  is in  $U$ , because  $U = G(U)$ , iff  $\#T(\lceil A \rceil)$  is in  $G(U)$  by rule (r2), iff  $T(\lceil A \rceil)$  is true, by Definition 3.1.
  - $\neg A$  is true in  $M$  iff  $\#[\neg A]$  is in  $G(U)$ , by rule (r1), iff  $\#[\neg A]$  is in  $U$ , because  $U = G(U)$ , iff  $\#[T(A)]$  is in  $F(U)$ , by rule (r2), iff  $T(\lceil A \rceil)$  is false, by Definition 3.1.
- Consequently, a sentence  $A$  of  $L$  is true in  $M$  iff  $T(\lceil A \rceil)$  is true, and false in  $M$  iff  $T(\lceil A \rceil)$  is false. Thus MTT proves the ZF set theory modelled by  $M$  true.
- (c)  $T$  is by Definition 3.1 a truth predicate for  $\mathcal{L}_U$ , and is not subject to any restrictions within  $\mathcal{L}_U$ .
  - (d) Let  $A$  be a sentence of  $\mathcal{L}_U$ . Since  $U$  is consistent, then  $\#A$  is by Lemma 2.1 either in  $G(U)$  or in  $F(U)$ . Applying rules (r2) and (r3), and the assumption  $U = G(U)$ , we obtain
    - $\#A$  is in  $G(U)$  iff  $\#A$  is in  $U$  iff  $\#T(\lceil A \rceil)$  is in  $G(U)$ ;
    - $\#A$  is in  $F(U)$  iff  $\#[\neg A]$  is in  $G(U)$  iff  $\#[\neg A]$  is in  $U$  iff  $\#T(\lceil A \rceil)$  is in  $F(U)$ .
 Consequently,  $\#A$  and  $\#T(\lceil A \rceil)$  are both either in  $G(U)$  or in  $F(U)$ . Thus  $\#[A \leftrightarrow T(\lceil A \rceil)]$  is by rule (r7) in  $G(U)$ , so that  $A \leftrightarrow T(\lceil A \rceil)$  is true by Definition 3.1. This holds for every sentence  $A$  of  $\mathcal{L}_U$ , whence  $T$ -biconditionals  $A \leftrightarrow T(\lceil A \rceil)$  are derivable unrestrictedly in  $\mathcal{L}_U$ .
  - (e) It follows from Definition 3.1 that ‘ $\#A$  is in  $G(U)$ ’ can be replaced by ‘ $A$  is true’ and ‘ $\#A$  is in  $F(U)$ ’ by ‘ $A$  is false’. Thus rules (r3)–(r9) imply that the truth in MTT is compositional.
  - (f) MTT allows for standard interpretations.  $M$  is by [16] a standard model of set theory.
  - (g) The outer logic and the inner logic coincide in MTT because of (d), as stated [19].
  - (h) By the proof of (e) the logic of  $\mathcal{L}_U$ , i.e., the inner logic of MTT is classical. This and (g) imply that also the outer logic of MTT is classical.
  - (i) Since  $U$  is consistent, it follows from Lemmas 2.1 and 2.2 that every sentence of  $\mathcal{L}_U$  is either true or false in the sense of Definition 3.1. Thus MTT is free from paradoxes. Only mathematical methods are used in its formulation, whence it is a mathematical theory.
  - (j) By Definition 3.1 truth for the language is defined within the language  $\mathcal{L}_U$  itself.  $\square$

*Remarks 3.1.* If  $\theta$  is a monadic predicate of  $L$ , there is such a sentence  $A$  in  $L$  that biconditionality  $A \leftrightarrow \neg\theta(\lceil A \rceil)$  is provable from axioms of ZF set theory (cf., e.g., [18, Lemma IV.5.31]). Thus  $T$ -biconditionality rule (d) does not hold in  $L$  when  $T = \theta$ . In particular,  $L$  does not contain its truth predicate. This holds also for the language  $\mathcal{L}$ . To satisfy all norms (a)–(j) it is then essential that those languages  $\mathcal{L}_U$  for which MTT is formulated differ both from  $L$  and from  $\mathcal{L}$ . This holds in our approach, since  $G(U) \cup F(U)$  is a proper and nonempty subset of  $D$  when  $U$  is a consistent fixed point of  $G$ .

The language  $\mathcal{L}_U$  corresponding to the smallest consistent fixed point  $U$  of  $G$  relates to the language of grounded sentences defined in [11, 17]. See also [6], where considerations are restricted to signed statements of  $\mathcal{L}$ .

MTT is only a *theory* of truth depending, e.g., on the semantic notion ‘true in  $M$ ’. Moreover, the construction of  $M$  starts from the set  $\omega \cup \{\omega\}$ , the existence of which is a matter of faith.

## 4 Appendix

Let  $\mathcal{P}$  denote the family of all consistent subsets of the set  $D$  of Gödel numbers of sentences of  $\mathcal{L}$ . Fixed points of the set mapping  $G := U \mapsto G(U)$  from  $\mathcal{P}$  to  $\mathcal{P}$ , i.e., those  $U \in \mathcal{P}$  for which  $U = G(U)$ , have a central role in the formulation of MTT. Before the proof of our main fixed point theorem we prove some auxiliary results.

**Lemma 4.1.** *Let  $U$  and  $V$  be sets of  $\mathcal{P}$ , and assume that  $U \subseteq V$ . If  $A$  is a sentence of  $\mathcal{L}$ , then  $\#A$  is in  $G(V)$  whenever it is in  $G(U)$ , and  $\#A$  is in  $F(V)$  whenever it is in  $F(U)$ .*

*Proof.* Assume that  $U \subseteq V$ . Consider first basic sentences. Let  $A$  be a sentence of  $L$ . By rule (r1)  $\#A$  is in  $G(U)$  and also in  $G(V)$  iff  $A$  is true in  $M$ , and both in  $F(U)$  and in  $F(V)$  iff  $\neg A$  is true in  $M$ .

Let  $A$  be a sentence of  $\mathcal{L}$ . If  $\#T(\lceil A \rceil)$  is in  $G(U)$ , then  $\#A$  is in  $U$  by rule (r2). Because  $U \subseteq V$ , then  $\#A$  belongs to  $V$ , whence  $\#T(\lceil A \rceil)$  is in  $G(V)$  by rule (r2).

If  $\#T(\lceil A \rceil)$  is in  $F(U)$ , then  $\#[\neg A]$  is in  $U$  by rule (r2). Since  $U \subseteq V$ , then  $\#[\neg A]$  belongs to  $V$ , so that  $\#T(\lceil A \rceil)$  is in  $F(V)$  by rule (r2).

Thus all basic sentences satisfy the lemma.

Assume that  $A$  is a sentence of  $\mathcal{L}$ . If  $\#[\neg A]$  is in  $G(U)$  but not in  $G(V)$ , then  $\#A$  is in  $F(U)$  but not in  $F(V)$  by rule (r3). If  $\#[\neg A]$  is in  $F(U)$  but not in  $F(V)$ , then  $\#A$  is in  $G(U)$  but not in  $G(V)$  by rule (r3). Hence, if  $A$  satisfies the lemma, then also  $\neg A$  satisfies it.

Make an induction hypothesis:

(h1)  $A$  and  $B$  are such sentences of  $\mathcal{L}$  that  $\#A$  is in  $G(V)$  if it is in  $G(U)$ ,  $\#A$  is in  $F(V)$  if it is in  $F(U)$ ,  $\#B$  is in  $G(V)$  if it is in  $G(U)$ , and  $\#B$  is in  $F(V)$  if it is in  $F(U)$ .

If  $\#[A \vee B]$  is in  $G(U)$ , then  $\#A$  or  $\#B$  is in  $G(U)$  by rule (r4). By (h1)  $\#A$  or  $\#B$  is in  $G(V)$ , so that  $\#[A \vee B]$  is in  $G(V)$ . If  $\#[A \vee B]$  is in  $F(U)$ , then  $\#A$  and  $\#B$  are in  $F(U)$  by rule (r4), and hence also in  $F(V)$ , by (h1), so that  $\#[A \vee B]$  is in  $F(V)$ . Thus  $A \vee B$  satisfies the lemma if (h1) holds.

If  $\#[A \wedge B]$  is in  $G(U)$ , then both  $\#A$  and  $\#B$  are in  $G(U)$  by rule (r5), and hence also in  $G(V)$ , by (h1). Thus  $\#[A \wedge B]$  is in  $G(V)$ . If  $\#[A \wedge B]$  is in  $F(U)$ , then  $\#A$  or  $\#B$  is in  $F(U)$  by rule (r5), and hence also in  $F(V)$ , by (h1), whence  $\#[A \wedge B]$  is in  $F(V)$ . Thus  $A \wedge B$  satisfies the lemma if (h1) holds.

If  $\#[A \rightarrow B]$  is in  $G(U)$ , then  $\#[\neg A]$  or  $\#B$  is in  $G(U)$ , i.e.,  $\#A$  is in  $F(U)$  or  $\#B$  is in  $G(U)$ . Then, by (h1),  $\#A$  is in  $F(V)$  or  $\#B$  is in  $G(V)$ , i.e.,  $\#[\neg A]$  or  $\#B$  is in  $G(V)$ . Thus  $\#[A \rightarrow B]$  is in  $G(V)$ . If  $\#[A \rightarrow B]$  is in  $F(U)$ , then  $\#A$  is in  $G(U)$  and  $\#B$  is in  $F(U)$ . This implies by (h1) that  $\#A$  is in  $G(V)$  and  $\#B$  is in  $F(V)$ , so that  $\#[A \rightarrow B]$  is in  $F(V)$ . Thus  $A \rightarrow B$  satisfies the lemma if (h1) holds.

If  $\#[A \leftrightarrow B]$  is in  $G(U)$ , then  $\#A$  and  $\#B$  are both in  $G(U)$  or both in  $F(U)$ , and hence both in  $G(V)$  or both in  $F(V)$  if (h1) holds, so that  $\#[A \leftrightarrow B]$  is in  $G(V)$ . If  $\#[A \leftrightarrow B]$  is in  $F(U)$ , then  $\#A$  is in  $G(U)$  and  $\#B$  is in  $F(U)$  or vice versa. Thus  $\#A$  is in  $G(V)$  and  $\#B$  is in  $F(V)$  or vice versa if (h1) holds, so that  $\#[A \leftrightarrow B]$  is in  $F(V)$ . Thus  $A \leftrightarrow B$  satisfies the lemma if (h1) holds. Because (h1) holds for basic sentences, the above results and induction on the complexity of expressions imply that logical connectives of sentences of  $\mathcal{L}$  satisfy the lemma.

It follows from the proof of Lemma 2.1 that  $\#[\exists xT(x)]$  is neither in  $F(U)$  nor in  $F(V)$ . If  $\#[\exists xT(x)]$  is in  $G(U)$ , then  $T(\mathbf{n})$  is in  $G(U)$  for some  $\mathbf{n} \in X$ . Since  $\mathbf{n} = \lceil A \rceil$  for some sentence  $A$  of  $\mathcal{L}$ , then  $\#T(\lceil A \rceil)$  is in  $G(U)$ . Thus  $\#A$  is by rule (r2) in  $U$ , and hence also in  $V$ . Then, by rule (r2),  $\#T(\lceil A \rceil) = \#T(\mathbf{n})$  is in  $G(V)$ , so that  $\#[\exists xT(x)]$  is in  $G(V)$  by rule (r8). Consequently,  $\#[\exists xT(x)]$  is in  $G(V)$  whenever it is in  $G(U)$ .

The proof of Lemma 2.1 implies that  $\#[\forall xT(x)]$  is neither in  $G(U)$  nor in  $G(V)$ . If  $\#[\forall xT(x)]$  is in  $F(U)$ , then  $T(\mathbf{n})$  is in  $F(U)$  for some  $\mathbf{n} \in X$ . But  $\mathbf{n} = \lceil A \rceil$  for some sentence  $A$  of  $\mathcal{L}$ , so that  $\#T(\lceil A \rceil)$  is in  $F(U)$ . So  $\#[\neg A]$  is by rule (r2) in  $U$ , and hence also in  $V$ . Thus, by rule (r2),  $\#T(\lceil A \rceil) = \#T(\mathbf{n})$  is in  $F(V)$ , so that  $\#[\exists xT(x)]$  is in  $F(V)$  by rule (r8). Consequently,  $\#[\forall xT(x)]$  is in  $F(V)$  whenever it is in  $F(U)$ .

The above results imply the conclusion of the lemma.  $\square$

According to Lemma 2.2 the mapping  $G := U \mapsto G(U)$  maps  $\mathcal{P}$  into  $\mathcal{P}$ . Assuming that  $\mathcal{P}$  is ordered by inclusion, the above lemma implies the following result.

**Lemma 4.2.**  *$G$  is order preserving in  $\mathcal{P}$ , i.e.,  $G(U) \subseteq G(V)$  whenever  $U$  and  $V$  are sets of  $\mathcal{P}$  and  $U \subseteq V$ .*

As above lemmas, also the next lemma has a counterpart in [11].

**Lemma 4.3.** (a) *If  $\mathcal{W}$  is a chain in  $\mathcal{P}$ , then the union  $\cup\mathcal{W} = \cup\{U \mid U \in \mathcal{W}\}$  is consistent.*  
(b) *The intersection  $\cap\mathcal{W} = \cap\{U \mid U \in \mathcal{W}\}$  of every nonempty subfamily  $\mathcal{W}$  of  $\mathcal{P}$  is a consistent subset of  $D$ .*

*Proof.* (a) Assume on the contrary that  $\cup\mathcal{W}$  is not consistent. Then there is a such a sentence  $A$  of  $\mathcal{L}$  that both  $\#A$  and  $\#[\neg A]$  are in  $\cup\mathcal{W}$ . Thus  $\mathcal{W}$  has a member, say  $U$ , which contains  $\#A$ , and a member, say  $V$ , which contains  $\#[\neg A]$ . If  $\mathcal{W}$  is a chain, then  $U \subseteq V$  or  $V \subseteq U$ . In former case  $V$  and in latter case  $U$  contains both  $\#A$  and  $\#[\neg A]$ . But this is impossible because  $\mathcal{W}$  is a subfamily of  $\mathcal{P}$ . This proves (a).

(b) The intersection  $\cap\mathcal{W}$  is a subset of  $D$ , and is contained in every member of  $\mathcal{W}$ . Hence  $\cap\mathcal{W}$  is consistent, for otherwise there is such a sentence  $A$  in  $\mathcal{L}$  that both  $\#A$  and  $\#[\neg A]$  are in  $\cap\mathcal{W}$ . Then every member of  $\mathcal{W}$  would also contain both  $\#A$  and  $\#[\neg A]$ . But this is impossible because every member of  $\mathcal{W}$  is consistent. This proves (b).  $\square$

In the formulation and the proof of our main fixed point theorem we use transfinite sequences of  $\mathcal{P}$  indexed by von Neumann ordinals. Such a sequence  $(U_\lambda)_{\lambda \in \alpha}$  of  $\mathcal{P}$  is said to be strictly increasing if  $U_\mu \subset U_\nu$  whenever  $\mu \in \nu \in \alpha$ , and strictly decreasing if  $U_\nu \subset U_\mu$  whenever  $\mu \in \nu \in \alpha$ . A set  $V$  of  $\mathcal{P}$  is called *sound* iff  $V \subseteq G(V)$ .

The following fixed point theorem is an application of Lemmas 2.2, 4.2 and 4.3.

**Theorem 4.1.** *If  $V \in \mathcal{P}$  is sound, then there exists the smallest of those fixed points of  $G$  which contain  $V$ . This fixed point is the last member of the union of those transfinite sequences  $(U_\lambda)_{\lambda \in \alpha}$  of  $\mathcal{P}$  which satisfy*

$$(C) \ (U_\lambda)_{\lambda \in \alpha} \text{ is strictly increasing, } U_0 = V, \text{ and if } 0 \in \mu \in \alpha, \text{ then } U_\mu = \bigcup_{\lambda \in \mu} G(U_\lambda).$$

*Proof.* Let  $V \in \mathcal{P}$  be sound. Transfinite sequences of  $\mathcal{P}$  satisfying (C) are called  $G$ -sequences. We shall first show that  $G$ -sequences are nested:

- (1) *Assume that  $(U_\lambda)_{\lambda \in \alpha}$  and  $(V_\lambda)_{\lambda \in \beta}$  are  $G$ -sequences, and that  $\{U_\lambda\}_{\lambda \in \alpha} \not\subseteq \{V_\lambda\}_{\lambda \in \beta}$ . Then  $(V_\lambda)_{\lambda \in \beta} = (U_\lambda)_{\lambda \in \beta}$ .*

By the assumption of (1)  $\mu = \min\{\lambda \in \alpha \mid U_\lambda \notin \{V_\lambda\}_{\lambda \in \beta}\}$  exists, and  $\{U_\lambda\}_{\lambda \in \mu} \subseteq \{V_\lambda\}_{\lambda \in \beta}$ . Properties (C) imply by transfinite induction that  $U_\lambda = V_\lambda$  for each  $\lambda \in \mu$ . To prove that  $\mu = \beta$ , make a counter-hypothesis:  $\mu \in \beta$ . Since  $\mu \in \alpha$  and  $U_\lambda = V_\lambda$  for each  $\lambda \in \mu$ , it follows from properties (C) that  $U_\mu = \bigcup_{\lambda \in \mu} G(U_\lambda) = \bigcup_{\lambda \in \mu} G(V_\lambda) = V_\mu$ , which is impossible, since  $V_\mu \in \{V_\lambda\}_{\lambda \in \beta}$ , but  $U_\mu \notin \{V_\lambda\}_{\lambda \in \beta}$ . Consequently,  $\mu = \beta$  and  $U_\lambda = V_\lambda$  for each  $\lambda \in \beta$ , whence  $(V_\lambda)_{\lambda \in \beta} = (U_\lambda)_{\lambda \in \beta}$ .

By definition, every  $G$ -sequence  $(U_\lambda)_{\lambda \in \alpha}$  is a function  $\lambda \mapsto U_\lambda$  from  $\alpha$  into  $\mathcal{P}$ . Property (1) implies that these functions are compatible. Thus their union is by [15, Theorem 2.3.12] a function with values in  $\mathcal{P}$ , the domain being the union of all index sets of  $G$ -sequences. Because these index sets are ordinals, then their union is also an ordinal by [18, I.8.10]. Denote it by  $\gamma$ . The union function can be represented as a sequence  $(U_\lambda)_{\lambda \in \gamma}$  of  $\mathcal{P}$ . It is strictly increasing as a union of strictly increasing nested sequences.

To show that  $\gamma$  is a successor, assume on the contrary that  $\gamma$  is a limit ordinal. Given  $\nu \in \gamma$ , then  $\mu = \nu \cup \{\nu\}$  and  $\alpha = \mu \cup \{\mu\}$  are in  $\gamma$ , and  $(U_\lambda)_{\lambda \in \alpha}$  is a  $G$ -sequence. Denote  $U_\gamma = \bigcup_{\lambda \in \gamma} G(U_\lambda)$ .  $G$  is order preserving by Lemma 4.2, and  $(U_\lambda)_{\lambda \in \gamma}$  is a strictly increasing sequence of  $\mathcal{P}$ . Thus  $\{G(U_\lambda)\}_{\lambda \in \gamma}$  is a chain in  $\mathcal{P}$ , whence  $U_\gamma$  is consistent by Lemma 4.3(a). Moreover,  $U_\nu \subset U_\mu = \bigcup_{\lambda \in \mu} G(U_\lambda) \subseteq U_\gamma$ . This holds for each  $\nu \in \gamma$ , whence  $(U_\lambda)_{\lambda \in \gamma \cup \{\gamma\}}$  is a  $G$ -sequence. This is impossible, since  $(U_\lambda)_{\lambda \in \gamma}$  is the union of all  $G$ -sequences. Consequently,  $\gamma$  is a successor, say  $\gamma = \alpha \cup \{\alpha\}$ . Thus  $U_\alpha$  is the last member of  $(U_\lambda)_{\lambda \in \gamma}$ ,  $U_\alpha = \max\{U_\lambda\}_{\lambda \in \gamma}$ , and  $G(U_\alpha) = \max\{G(U_\lambda)\}_{\lambda \in \gamma}$ . Moreover,  $(U_\lambda)_{\lambda \in \gamma}$  is a  $G$ -sequence, for otherwise  $(U_\lambda)_{\lambda \in \alpha}$

would be the union of all  $G$ -sequences. In particular,  $U_\alpha = \bigcup_{\lambda \in \alpha} G(U_\lambda) \subseteq \bigcup_{\lambda \in \gamma} G(U_\lambda) = G(U_\alpha)$ , so that  $U_\alpha \subseteq G(U_\alpha)$ . Equality holds, since otherwise the longest  $G$ -sequence  $(U_\lambda)_{\lambda \in \gamma}$  could be extended by  $U_\gamma = \bigcup_{\lambda \in \gamma} G(U_\lambda)$ . Thus  $U_\alpha$  is a fixed point of  $G$  in  $\mathcal{P}$ .

Assume that  $W \in \mathcal{P}$  is a fixed point of  $G$ , and that  $V \subseteq W$ . Then  $U_0 = V \subseteq W$ . If  $0 \in \mu \in \gamma$ , and  $U_\lambda \subseteq W$  for each  $\lambda \in \mu$ , then  $G(U_\lambda) \subseteq G(W)$  for each  $\lambda \in \mu$ , whence  $U_\mu = \bigcup_{\lambda \in \mu} G(U_\lambda) \subseteq G(W) = W$ . Thus, by transfinite induction,  $U_\mu \subseteq W$  for each  $\mu \in \gamma$ . Consequently,  $U_\alpha \subseteq W$  so that  $U_\alpha$  is the smallest fixed point of  $G$  that contains  $V$ .  $\square$

As a consequence of Theorem 4.1 we obtain.

**Corollary 4.1.**  *$G$  has the smallest fixed point.*

*Proof.* The empty set  $\emptyset$  is both sound and consistent. Thus the smallest fixed point of  $G$  that contains  $\emptyset$  exists by Theorem 4.1. It is the smallest fixed point of  $G$ , since every fixed point of  $G$  contains  $\emptyset$ .  $\square$

Next we prove that every consistent subset of  $D$  has the greatest sound and consistent subset.

**Theorem 4.2.** *The equation  $V = W \cap G(V)$  has for each consistent subset  $W$  of  $D$  the greatest solution  $V$  in  $\mathcal{P}$ . It is the greatest sound set of  $\mathcal{P}$  that is contained in  $W$ .  $V$  is the last member of the union of those transfinite sequences  $(V_\lambda)_{\lambda \in \alpha}$  of  $\mathcal{P}$  which satisfy*

$$(D) \ (V_\lambda)_{\lambda \in \alpha} \text{ is strictly decreasing, } V_0 = W, \text{ and if } 0 \in \mu \in \alpha, \text{ then } V_\mu = W \cap \left( \bigcap_{\lambda \in \mu} G(V_\lambda) \right).$$

*Proof.* Assume that  $W$  is a consistent subset of  $D$ . Like the proof that the union of all  $G$ -sequences is a  $G$ -sequence one can prove that the union of all transfinite sequences which have properties (D) given in Theorem 4.2 has property (D). Let  $(V_\lambda)_{\lambda \in \gamma}$  be that sequence. Denote  $V = W \cap \left( \bigcap_{\lambda \in \gamma} G(V_\lambda) \right)$ . Because  $W$  and the sets  $V_\lambda$ ,  $\lambda \in \gamma$ , are consistent, it follows from Lemma 2.2 and Lemma 4.3(b) that  $V$  is consistent. Moreover,  $V \subseteq V_\lambda$  for each  $\lambda \in \gamma$ . If  $V \subset V_\lambda$  for each  $\lambda \in \gamma$ , then the choice  $V_\lambda = V$  implies that  $(V_\lambda)_{\lambda \in \gamma \cup \{\gamma\}}$  satisfies (D) when  $\alpha = \gamma \cup \{\gamma\}$ . But this is impossible because of the choice of  $(V_\lambda)_{\lambda \in \gamma}$ . Thus  $V = \min\{V_\lambda\}_{\lambda \in \gamma}$ , and  $V$  is the last member of  $(V_\lambda)_{\lambda \in \gamma}$  because this sequence is strictly decreasing. Since  $G$  is order preserving, then  $G(V) = \min\{G(V_\lambda)\}_{\lambda \in \gamma} = \bigcap_{\lambda \in \gamma} G(V_\lambda)$ . Thus  $V = W \cap G(V)$ , so that

$V \subseteq G(V)$ , i.e.,  $V$  is sound and is contained in  $W$ .

Assume that  $U$  is consistent, that  $U \subseteq G(U)$ , and that  $U \subseteq W$ . Since  $V_0 = W$  by (D), then  $U \subseteq V_0$ . If  $0 \in \mu \in \gamma$  and  $U \subseteq V_\lambda$  for each  $\lambda \in \mu$ , then  $G(U) \subseteq G(V_\lambda)$  for each  $\lambda \in \mu$ , whence  $U \subseteq W \cap G(U) \subseteq W \cap \left( \bigcap_{\lambda \in \mu} G(V_\lambda) \right) = V_\mu$ . Thus, by transfinite induction,  $U \subseteq V_\lambda$  for each

$\lambda \in \gamma$ , so that  $U \subseteq \min\{V_\lambda\}_{\lambda \in \gamma} = V$ . Consequently,  $V$  is the greatest sound and consistent subset of  $D$  that is contained in  $W$ .  $\square$

The following results shows that if  $W$  is any consistent set of Gödel numbers of sentences of  $L$ , and  $V$  is determined by Theorem 4.2, then the union of the  $G$ -sequences (C) of Theorem 4.1 is the smallest fixed point of  $G$ .

**Corollary 4.2.** *Let  $W$  be a consistent subset of  $D$  containing only Gödel numbers of sentences of  $L$ . Then*

- (a)  $V = W \cap G(W)$  is the greatest consistent solution of  $V = V \cap G(V)$  contained in  $W$ .
- (b) The union of the  $G$ -sequences (C) of  $V = W \cap G(W)$  is the smallest fixed point of  $G$ .

*Proof.* (a) Let  $W$  be a consistent set of Gödel numbers of sentences of  $L$ .  $G(W)$  contains by rule (r1) Gödel numbers of those sentences of  $L$  which are true in  $M$ , and only those. Thus  $V = W \cap G(W)$  is the set of Gödel numbers of those sentences of  $W$  which are true in  $M$ . Hence  $V \subseteq G(\emptyset) \subseteq G(V)$ , so that  $V$  is sound. It is also consistent, as an intersection of two consistent sets. Consequently,  $V = W \cap G(W)$  is a solution of  $V = V \cap G(V)$ , and by Theorem 4.2 the greatest solution contained in  $W$ .

(b)  $V = W \cap G(W)$  is by (a) sound and consistent, and  $V \subseteq G(\emptyset)$ . If  $U$  is a fixed point of  $G$ , then  $V \subseteq G(\emptyset) \subset G(U) = U$ . Thus  $V$  is contained in the smallest fixed point of  $G$ . By Theorem 4.1, the union of the  $G$ -sequences (C) is the smallest fixed point of  $G$  that contains  $V$ . This proves (b).  $\square$

*Remarks 4.1.* The smallest members of  $(U_\lambda)_{\lambda \in \alpha}$  satisfying (C) are  $n$ -fold iterations  $U_n = G^n(V)$ ,  $n \in \mathbb{N} = \{0, 1, \dots\}$ . If they form a strictly increasing sequence, the next member  $U_\omega$  is their union,  $U_{\omega+n} = G^n(U_\omega)$ ,  $n \in \mathbb{N}$ , and so on.

If the set  $W$  is finite, then the longest sequence  $(V_\lambda)_{\lambda \in \alpha}$  satisfying (D) is obtained by the finite algorithm:  $V_0 = W$ . For  $n$  from 0 while  $V_n \neq W \cap G(V_n)$  do:  $V_{n+1} = W \cap G(V_n)$ .

Zorn's Lemma, together with Lemmas 4.2 and 4.3, can be applied, e.g., as in [6, 11], to prove results of Theorem 4.1 and Corollary 4.1, and also the existence of maximal fixed points of  $G$ . ZF set theory is no more sufficient framework for such proofs, since Zorn's Lemma is equivalent to the Axiom of Choice. The result of Theorem 4.2 that gives a method to determine sound subsets of  $D$  seems to be new.

As for generalizations of Theorems 4.1 and 4.2, see, e.g., [2].

**Acknowledgments:** The author is indebted to Prof. Hannes Leitgeb and to Ph.D. Tapani Hyttinen for valuable discussions on the subject. The present work is influenced by [11].

## References

- [1] Beall, Jc & Glandzberg, Michael (2013) Liar Paradox, The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.). Routledge.
- [2] Carl, S. & Heikkilä, S. (2011) Fixed Point Theory in Ordered Sets and Applications, Springer, Berlin Heidelberg New York. theory, Bulletin of the American Mathematical Society, **69**, 537–540.
- [3] Cohen, Paul (1963) A minimal model for set theory, Bulletin of the American Mathematical Society, **69**, 537–540.
- [4] Feferman, Solomon (1984) Toward useful type-free theories, Journal of Symbolic Logic, **49**, 1, 75–111.
- [5] Feferman, Solomon (2012) Axiomatizing truth. Why and how? Logic, Construction, Computation (U. Berger et al. eds.) Ontos Verlag, Frankfurt, 185-200.
- [6] Fitting, Melvin (1986) Notes on the Mathematical Aspects of Kripke's Theory of Truth, Notre Dame Journal of Formal Logic, **27**,1, 75–88.
- [7] Fitting, Melvin (2007), Incompleteness in the Land of Sets, Studies in Logic.
- [8] Halbach, Volker (1997) Tarskian and Kripkean truth, Journal of Philosophical Logic, **26**, 1, 69–80.
- [9] Halbach, Volker (2011) Axiomatic Theories of Truth, Cambridge University Press.
- [10] Halbach, Volker & Horsten, Leon (2006) Axiomatizing Kripke's Theory of Truth, J, Symbolic Logic **71**,2, 677–712.
- [11] Heck, Richard G. , Kripke's theory of truth, <http://www.frege.brown.edu/heck/philosophy/pdf/notes/KripkesTheoryOfTruth.pdf>.
- [12] Heck, Richard G. Formal background for theories of truth, <http://www.frege.brown.edu/heck/philosophy/pdf/notes/FormalBackground.pdf>.
- [13] Heikkilä, S. (2014) A Mathematical Theory of Truth with Applications, arXiv.1307.4692.
- [14] Hodges, Wilfried (2013) Tarski's Truth Definitions, The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.).
- [15] Hrbacek, Karel & Jech, Thomas (1999) Introduction to Set Theory, Marcel Dekker, Inc, New York - Basel.
- [16] Kanamori, Akihiro (2008) Cohen and Set Theory, The Bulletin of Symbolic Logic, **14**,3, 351–377.
- [17] Kripke, Saul (1975) Outline of a Theory of Truth, Journal of Philosophy, **72**, 690–716.
- [18] Kunen, Kenneth (2009) The Foundations of Mathematics, Mathematical Logic and Foundations, **19**.
- [19] Leitgeb, Hannes (2007) What Theories of Truth Should be Like (but Cannot be), Philosophy Compass, **2**/2, 276–290.
- [20] McGee, Vann (1989) Applying Kripke's Theory of Truth, The Journal of Philosophy, **86**,10, 530–539.
- [21] Tarski, A (1955) A lattice-theoretical fixpoint theorem and its applications, Pacific T. Math. **5**, 285–309.
- [22] Tarski, Alfred (1936) Der Wahrheitsbegriff in Formalisierten Sprachen, Studia Philosophica 1: 261–405.