Logique & Analyse **185–188** (2004), 463–468

A NOTE CONCERNING THE NOTION OF SATISFIABILITY

LEON HORSTEN

Abstract
Tarski has shown how the argumentation of the liar paradox can be used to prove a theorem about truth in formalized languages. In this paper, it is shown how the paradox concerning the least undefinable ordinal can be used to prove a no go-theorem concerning the notion of satisfaction in formalized languages. Also, the connection of this theorem with the absolute notion of definability is discussed.

Julius König at one point claimed to have proved that the real numbers (R) cannot be wellordered.[1] He argued more or less as follows. Suppose there were a wellordering $<$ of R. R is nondenumerable, while there are only denumerably many definitions of real numbers. Therefore there must be (uncountably many) undefinable real numbers. By the properties of $<$, there must be a unique least undefinable real number. Call this number r. But r is defined by the clause "the least undefinable real number". Contradiction.

The set theoretic community has never accepted König's argument. First, the set theoretic community holds that the absolute notion of definability is not a definite mathematical notion. The occurrences of 'definability' in the argument should, in its opinion, be replaced by 'definability in L', for some formal language L which the proponent of the argument is allowed to choose freely.[2] Definability in a formal language is a perfectly determinate mathematical concept. But if one makes the suggested substitution, König's argument no longer goes through as it stands. For one would now have to show that definability in L is itself expressible in L. With the benefit of hindsight, we know this to be a tall order. Second, König implicitly assumes that the hypothetical wellordering of the real numbers is definable. This assumption would also have to be established for the argument to be persuasive. Again, the history of set theory has born out that this is no easy task.

[1] See [KO].

[2] See [ZE, p. 192, fn. 11].

Alan Hazen has constructed an argument concerning the definability of ordinal numbers which uses König's reasoning.[3] Hazen's simple argument purports to show that every ordinal number is definable. It goes roughly as follows. Suppose, for a reductio, that there is an undefinable ordinal. Then by the definable wellordering of the ordinals, there is a least undefinable ordinal o. But then we have just given a definition of this ordinal, so o is definable after all. Contradiction.

The second objection by the mathematical community then no longer applies. For a wellordering of the ordinals is definable: on the usual way of defining the ordinals, the $<$-relation wellorders them. The first objection by the mathematicical community does apply to Hazen's argument. Mathematicians will insist that the notion of definability be made precise as definability in some formal language L. Moreover, for the argument to go through, this language L must be the language in which Hazen's argument is formulated.

The relevant notion of definability of objects can be defined in terms of the notion of satisfaction ('true of'): an object x is definable if and only if there is a predicate which is true of x and of nothing else. Let L_S be the language of first-order set theory plus a primitive satisfaction predicate $Sat(x, y)$. Then 'x is definable' can be expressed in L_S as:

$$\exists y[FormulaL_S(y) \wedge Sat(x, y) \wedge \forall z(z \neq x \rightarrow \neg Sat(z, y))],$$

where $FormulaL_S(x)$ expresses in L_S that x is the name (gödel number) of a formula of L_S. Let us abbreviate this formula as $Def(x)$.

We have seen that Hazen's informal argument does not immediately carry conviction as it stands. But the foregoing considerations do open up the question whether the reasoning of Hazen's (and of König's) argument might not be used to prove a proposition that is of philosophical interest. After all, Tarski has done something similar for another infamous argument. Tarski has shown us how the reasoning of the liar paradox can be used to prove a theorem about truth in formalized languages,[4] namely (roughly) that no formalized language can contain its own truth predicate.

Consider the theory $ZFCSat$, which is formulated in the language L_S. $ZFCSat$ is defined as consisting of:

(1) the logical and set-theoretic axioms schemes and rules of ZFC ranging over the entire language L_S.[5]

[3] [HA, p. 18–19].

[4] See [TA].

[5] This is essential for what follows. If the logical and set-theoretic schemes were taken to range only over the fragment of the language not containing Sat, then the argument concerning the notion of satisfaction that I am about to formulate would not go through.

(2) a rule of inference for Def, which will be called $Rdef$:

$$\frac{\exists! x\phi(x)\ \ (with\ \phi\ containing\ only\ x\ free)}{\forall x(\phi(x) \rightarrow Def(x))}$$

In this rule, ϕ again ranges over the entire language L_S.

The restriction in (2) that ϕ is not allowed to contain free occurrences of variables other than x is essential here. For suppose ϕ would also contain another variable y free. Then in the premise of the rule, y could be regarded as implicitly universally quantified over. But the premise would not guarantee, for every y, the existence of a name for the x such that $\phi(x, y)$. For the expression "the x such that ϕ" would not then be a closed term, and hence would not be a complete (complex) name. To obtain the needed closed term, a name of y would also be needed.

With the restriction in place, $RDef$ seems a perfectly legitimate rule of inference. Nevertheless, $ZFCSat$ is inconsistent. The proof goes as follows. ZFC proves that there are only denumerably many formulas of L_S. Let $Ord(x)$ express in the language of ZFC that x is an ordinal. ZFC proves that there are nondenumerably many ordinals. So

$$ZFCSat \vdash \exists x(Ord(x) \wedge \neg Def(x))$$

Let $x < y$ express (in L_S): 'x is a smaller ordinal than y'. Then by the provable wellordering of the ordinals, ZFC proves:

$$\exists! x[Ord(x) \wedge \neg Def(x) \wedge \forall y[(Ord(y) \wedge y < x) \rightarrow Def(y)]]$$

Abbreviate the part in square brackets as $\theta(x)$. Then $ZFCSat \vdash \exists! x\theta(x)$. So by $RDef$, we have $ZFCSat \vdash \forall x(\theta(x) \rightarrow Def(x))$. But by sheer logic we also have $ZFCSat \vdash \forall x(\theta(x) \rightarrow \neg Def(x))$. This gives us a contradiction.

Although the proof is elementary,[6] I find this a remarkable result in the field of axiomatic approaches to the semantic paradoxes. The introduction rule for Def seems a very weak axiomatic theory of satisfaction for ZFC. Essentially, it only postulates an analogue for the definability predicate Def

[6] The proof is related to a known problem concerning the addition of axioms governing the description operator to systems of quantified modal logic. See [CA, p. 184], and more explicitly [RE, p. 452–453]. This connection will not be pursued further in this paper.

of the *Necessitation Rule*

$$\frac{\phi}{\Box\phi}$$

of modal logic. $RDef$ seems to say nothing about how satisfaction interacts with the logical connectives, or about how a closed sentence ϕ being true (i.e. satisfied by every x) implies ϕ. Appearances deceive.

It is clear that the above argument would go through against the background of a base theory much weaker than ZFC. Roughly, it suffices to have a base theory which can code its own syntax and which postulates a nondenumerable collection of objects with a definable wellordering on it. So we can informally phrase our findings as follows: no consistent and sufficiently strong theory contains the rule which infers from the statement that there is exactly one object for which the property ϕ holds to the conclusion that the object satisfying ϕ is definable.

We cannot consistently have the class of all unrestricted instances of the Tarski-biconditionals as our theory of truth.[7] In fact, a strenthening of the argument of the *Paradox of the Knower*[8] shows that already if the axiom schemes

$$T\ulcorner\varphi\urcorner \rightarrow \phi \tag{3}$$

$$T\ulcorner T\ulcorner\phi\urcorner \rightarrow \phi\urcorner \tag{4}$$

are added to PA (where the truth predicate T is allowed in the induction scheme),[9] a contradiction arises.[10] However, it *is* consistent to add the axiom scheme $T\ulcorner\varphi\urcorner \rightarrow \phi$ to PA, and it *is* consistent to add the analogue

$$\frac{\phi}{T\ulcorner\phi\urcorner}$$

for the truth predicate of the necessitation rule to PA (again, with T allowed in the induction scheme). In fact, the resulting theories are easily seen to be arithmetically conservative over PA. The Paradox of the Knower and its relatives show that we need little more than the predicate analogue of

[7] See [TA].

[8] See [KM].

[9] We use $\ulcorner\phi\urcorner$ to refer to the gödel number of ϕ.

[10] See [CR, p. 324].

the *Reflexivity Axiom* $\Box\varphi \rightarrow \varphi$ of modal logic to generate a paradox. Our argument shows that we need little more than the predicate analogue of the rule-version of the *converse* direction of the Tarski-biconditionals, i.e., the predicate analogue of the necessitation rule, in order to yield a contradiction.

Axiomatic approaches to truth and the semantic approaches have been gaining influence in recent years, both in the logical and in the philosophical literature.[11] It has emerged that there exist natural consistent formal theories of truth which capture many of the properties which we would naively be inclined to ascribe to truth. The above considerations provide support for the hypothesis that in the context of a stronger mathematical background theory and for the more general notion of satisfaction, the situation is in a sense less satisfactory: more seemingly fundamental properties of this semantical notion will have to be left out or at least restricted in any consistent axiomatic formalization.

Our argument is a proposition about the notion of satisfaction in formalized languages. It is less clear whether this argument can be regarded as an argument about *the absolute notion of definability*.

Hazen calls an object (absolutely) definable if and only if it is definable in some humanly usable language.[12] Hazen argues in his unpublished paper that there can be at most denumerably many possible languages in this sense, and that each such language is itself denumerable. Thus there exist only denumerably many 'possible concepts'.[13] This implies that his informal argument appears to be valid. And it would seem, then, that a version of our formal argument concerning the notion of satisfaction can be carried out for the union of all humanly possible languages. So Hazen has his work cut out for him. His paper consists largely in an ingenious and subtle attempt to find a way of dodging the conclusion of his informal argument.

Someone might agree with Hazen's characterization of absolute definability as definability in some humanly usable language but hold that there are nondenumerably many possible languages, and that there may therefore well be nondenumerably many definable objects. On this view, Hazen's informal argument may simply be sound. Also, our formal argument about the notion of satisfaction could not then be transformed in any obvious way into a no-go argument about the absolute notion of definability. The complaint would be that Def is just too narrow a characterization of the absolute notion of

[11] [CT] is a standard reference work on axiomatic theories of truth.

[12] [HA, p. 10].

[13] [HA, p. 18].

"24horsten"
2005/7/18
page 468

definability. In fact, the argument might then be taken to show that there must be a *proper class* of definable objects.

University of Leuven
Kardinaal Mercierplein 2
3000 Leuven, Belgium
E-mail: Leon.Horsten@hiw.kuleuven.be

REFERENCES

[CT] Cantini, A. *Logical Frameworks for Truth and Abstraction. An axiomatic study.* Elsevier, Amsterdam, 1996.

[CA] Carnap, R. *Meaning and Necessity. A study in semantics and modal logic.* Second Edition, University of Chicago Press, 1956.

[CR] Cross, Ch. *The paradox of the knower without epistemic closure.* Mind Vol. 110 No. 438 (April 2001), p. 319–333.

[HA] Hazen, A. *The contradiction of the least indefinable ordinal.*Typewritten manuscript, January 1982, 76 p.

[KO] König, J. *On the foundations of set theory and the continuum problem.*(1905) Reprinted in: J. Van Heijenoort (ed.) *From Frege to Gödel. A source book in mathematical logic 1879–1931.* Harvard University Press, 1967, p. 145–149.

[KM] Kaplan, D. and Montague, R. *A paradox regained.* Notre Dame Journal of Formal Logic Vol. 1 No 3(1960), p. 79–90.

[RE] Reinhardt, W. *Epistemic theories and the interpretation of Gödel's incompleteness theorems.* Journal of Philosophical Logic 15 (1986), p. 427–474.

[TA] Tarski, A. *The concept of truth in formalised languages.* (1935), Reprinted in: Tarski, A. *Logic, Semantics, Metamathematics. Papers from 1923 to 1938.* Second Edition, Hackett, 1983, p. 152–278.

[ZE] Zermelo, E. *A new proof of the possibility of a well-ordering.* (1908) Reprinted in: J. Van Heijenoort (ed.) *From Frege to Gödel. A source book in mathematical logic 1879–1931.* Harvard University Press, 1967, p. 183–198.