

Truth and existence

Jan Heylen

Leon Horsten

Abstract

Halbach has argued that Tarski biconditionals are not ontologically conservative over classical logic, but his argument is undermined by the fact that he cannot include a theory of arithmetic, which functions as a theory of syntax. This article is an improvement on Halbach's argument. By adding the Tarski biconditionals to inclusive negative free logic and the universal closure of minimal arithmetic, which is by itself an ontologically neutral combination, one can prove that at least one thing exists. The result can then be strengthened to the conclusion that infinitely many things exist. Those things are not just all Gödel codes of sentences but rather all natural numbers. Against this background inclusive negative free logic collapses into non-inclusive free logic, which collapses into classical logic. The consequences for ontological deflationism with respect to truth are discussed.

Keywords Disquotational theory of truth; Tarski biconditionals; ontological deflationism; inclusive negative free logic; universal closure of arithmetic; nominalism; finitism

Introduction

In this article we will focus primarily on the so-called *disquotational theory of truth* (DT), which goes back at least to Tarski (1936). Let \mathcal{L}_A be the language of arithmetic (Boolos et al., 2003, p. 103), which we use because syntax can be arithmetized (Boolos et al., 2003, ch 15). Furthermore, we will make use of the axioms of minimal arithmetic (M),¹ because of the representability of syntax in it (Boolos et al., 2003, ch. 16). Let $\ulcorner \phi \urcorner$ be the numeral denoting the Gödel number of sentence ϕ (Boolos et al., 2003, p. 221). Moreover, let T be the truth predicate. Then DT contains, first, classical logic (CL), second, the axioms of M (Boolos et al., 2003, p. 208) and, third, the Tarski-biconditionals:

$$T\ulcorner \phi \urcorner \leftrightarrow \phi, \tag{1}$$

for all $\phi \in \mathcal{L}_A$. (The restriction has to do with the Liar paradox.)

Tarski (1936) himself was of the opinion that DT is too weak, but he stipulated that only theories of truth that imply DT are materially adequate. Others thought that DT is adequate on its own. The latter point of view is related in part to a negative thesis about truth: any theory of truth has to be *deflationary*. The basic idea is that a truth theory is deflationary only if it does

¹The reader should be warned that Boolos et al. (2003, p. 207) use the label 'Q' for minimal arithmetic, whereas 'Q' is commonly used for Robinson arithmetic, which Boolos et al. (2003, p. 215) label as 'R'.

not have philosophical consequences.² In a famous paper (Tarski, 1944, p. 362) wrote about his semantic conception of truth being free of *epistemological* consequences:

Thus, we may accept the semantic conception of truth without giving up any epistemological attitude we may have had; we may remain naive realists, critical realists or idealists, empiricists or metaphysicians – whatever we were before. The semantic conception is completely neutral toward all these issues.

Here we are mainly interested in whether a theory of truth has *ontological* consequences.

An example of a theory of truth that is not ontologically deflationary is the correspondence theory of truth, which states that a proposition (or interpreted sentence or judgement) is true if and only if it corresponds to a fact. In consequence, this theory comes with an ontology of facts, even structured facts (e.g., negative facts or disjunctive facts). In contrast, note that DT does not mention facts, let alone structured facts. In this article we will explore whether DT is really ontologically deflationary.

It was a major contribution of Tarski to provide us with formally precise truth theories. The notion of a deflationary truth theory itself is not formally precise though. One way of making it formally precise is by wedding it to the notion of a deductively conservative theory.³ Theory T_2 is an extension of a theory T_1 if and only if every theorem of T_1 is a theorem of T_2 . Furthermore, theory T_2 is a deductively conservative extension of a theory T_1 if and only if

1. T_2 is an extension of T_1 ;
2. every theorem of T_2 belonging to the language of T_1 is also a theorem of T_1 .

In case T_2 is a truth theory and T_1 is any philosophical (e.g., epistemological, ontological, semantic) theory, then T_2 is a deflationary truth theory only if it is deductively conservative over T_1 .

In this article we will focus on whether or not a truth theory is ontologically conservative. This notion can also be made to a certain extent more precise as follows. Theory T_2 is an ontologically conservative extension of a theory T_1 only if every theorem of T_2 that says that there are at least n things is also a theorem of T_1 . Note that a deductively conservative extension of a theory is also an ontologically conservative extension of that theory, but not necessarily the other way around.

It should be remarked that the fact that a certain theory is not ontologically conservative over a base theory does neither tell us *how many beings* one gets in addition to the ones one already had through the base theory (if any) nor does it tell us what *kinds of beings* one gets over and above the ones belonging to the ontology of the base theory (if any). These further issue should be addressed as well if one wants to evaluate how deflationary a theory of truth is.

In this article we are interested in whether (a version of) DT is ontologically conservative over a system of inclusive free logic.⁴ In inclusive free logic domains of quantification can be

²Some go further and demand that a truth theory should not have mathematical or scientific consequences either (Horwich, 1998).

³This has been discussed in (Horsten, 1995; Shapiro, 1998; Ketland, 1999; Field, 1999; Horsten, 2011).

⁴See Nolt (2014) for an introduction to free logic.

empty (inclusive logic) and singular terms can be non-denoting (free logic), which contrasts with classical logic on both scores. One reason for our interest in inclusive free logic is that it tightly connects the existential commitment of a theory to its existential theorems (Bricker, 2014, Section 1.6.1). Another reason is that it is itself an ontologically neutral logic, which makes it an ideal base theory for our purposes.

There are different kinds of free logic: in negative free logic (NFL) atomic sentences with non-denoting terms are false, whereas in positive free logic they can be true and in neutral free logic they are without truth-value. In this article we are going to focus on NFL. The reason for this will become clear later. Before we will look into that let us see whether (a version of) DT is ontologically conservative over CL.

1 The disquotational theory of truth and classical logic

The existential commitment of CL is well-known. It is a theorem that at least one thing exists. The question is whether a theory of truth should be conservative over CL. According to Halbach (2001, p. 170) some philosophers are of the opinion that it should:

What seems to be of primary importance to Shapiro (1998) and Field (1999) is conservativeness over logic, i.e., over the empty base theory, or, equivalently, conservativeness over arbitrary, possibly empty, base theories.

Of course, since DT contains M, it is not ontologically conservative over CL by default. But Halbach (2001, p. 178-180) considered a version of DT, called TE, which contains only the Tarski biconditionals (1) but not M. He then showed that TE is not ontologically conservative over classical logic and, hence, it is not deductively conservative.

Two instances of scheme (1) are the following:

$$T \ulcorner \phi \vee \neg \phi \urcorner \leftrightarrow (\phi \vee \neg \phi); \quad (2)$$

$$T \ulcorner \phi \wedge \neg \phi \urcorner \leftrightarrow (\phi \wedge \neg \phi). \quad (3)$$

The following is an instantiation of the principle of the substitutivity of identicals:

$$\ulcorner \phi \vee \neg \phi \urcorner = \ulcorner \phi \wedge \neg \phi \urcorner \rightarrow (T \ulcorner \phi \vee \neg \phi \urcorner \rightarrow T \ulcorner \phi \wedge \neg \phi \urcorner). \quad (4)$$

It then follows tautologically from schemes (2), (3) and (4) that:

$$\ulcorner \phi \vee \neg \phi \urcorner \neq \ulcorner \phi \wedge \neg \phi \urcorner. \quad (5)$$

Existential generalisation yields

$$\exists x \exists y (x \neq y).$$

One cannot prove this in classical first-order logic with identity alone: just consider a model with a singleton as its domain of quantification. So, TE is not ontologically conservative over classical logic with identity. Therefore, it is also not deductively conservative.

Halbach's argument leads to an interesting conclusion, but Horsten (2011, p. 82) has criticized the idea that TE is a theory *of truth*:⁵

⁵A similar worry about Halbach's argument has been expressed by an anonymous reviewer.

[I]n the absence of an arithmetical theory functioning as a theory of syntax, it is not even clear what a truth theory such as [TE] *means*. One would expect that in the axiomatic frame of mind, not only the meaning of the truth predicate but also the meaning of the syntax theory is given by axioms describing syntactical operations. In the setting we work in, this role is played by Peano arithmetic.

If this is right, then Halbach should consider DT rather than TE, but DT is deductively conservative over M (Horsten, 2011, p. 75) and, hence, it passes the test for ontologically conservativeness as well. In what follows we will try to improve on this score. Surprisingly, we will do so by turning our attention to NFL.

2 The disquotational theory of truth and negative free logic

The language of NFL is the same as the language of CL. Let us use the following definition:

$$E!t =_{df} \exists x(x = t),$$

for any term t . Apart from the axiom schemes of sentential logic, the following are the axiom schemes of inclusive NFL:

A1 $\phi \rightarrow \forall x\phi$, with x not free in ϕ ;

A2 $\forall x(\phi \rightarrow \psi) \rightarrow (\forall x\phi \rightarrow \forall x\psi)$

A3 $\forall x\phi$, if $\phi(t/x)$ is an axiom;

A4 $\forall x\phi \rightarrow (E!t \rightarrow \phi(t/x))$;

A5 $\forall x(x = x)$;

A6 $t = t' \rightarrow (\phi \rightarrow \phi')$, with ϕ' identical to ϕ except that zero or more occurrences of t have been replaced by t' ;

A7 $P(t_1, \dots, t_n) \rightarrow E!t_i$, with $1 \leq i \leq n$ and with P any n -place predicate, including the identity predicate;

A8 $E!f(t_1, \dots, t_n) \rightarrow E!t_i$, with $1 \leq i \leq n$.

The only rule of inference is modus ponens. Axiom scheme **A4** is characteristic for *free* logic. Its classical counterpart is:

A4c $\forall x\phi \rightarrow \rightarrow \phi(t/x)$.

Axiom schemes **A7** and **A8** are characteristic for *negative* free logic. The absence of $\exists xE!x$ in the axiomatic base makes this an *inclusive* NFL.

In addition to inclusive NFL we are going to use the universal closure of the axioms of M:

M1 $\forall x\neg(x' = \mathbf{0})$;

M2 $\forall x \forall y (x' = y' \rightarrow x = y)$;

M3 $\forall x (x + \mathbf{0} = x)$;

M4 $\forall x \forall y (x + y' = (x + y)')$;

M5 $\forall x (x \cdot \mathbf{0} = \mathbf{0})$;

M6 $\forall x \forall y (x \cdot y' = (x \cdot y) + x)$.

M7 $\forall x \neg x < \mathbf{0}$.

M8 $\forall x \forall y (x < y' \leftrightarrow (x < y \vee x = y))$.

M9 $\forall x \forall y (x < y \vee x = y \vee y < x)$.

Let us call this theory M^\forall . As emphasized in (Boolos et al., 2003, p. 207-208) the axioms of M itself are the universal closures of the open formulas, but there is a convention to write down only the open formulas. In other words, M^\forall is just the official version of M but not the conventional version. The reason for using M^\forall is twofold. First, the extension of inclusive NFL with M^\forall is ontologically conservative, because universally quantified sentences are trivially true in models with an empty domain. Second, M^\forall is able to prove any sentence conventional M is able to prove if the relevant existence conditions are met. Indeed, based on **A4** one can derive, for instance, $E!t \rightarrow \neg(t' = \mathbf{0})$ and likewise for the other axioms. This means that in effect M^\forall is a *conditional syntax theory*: all the relevant facts about syntax can be represented in M^\forall , as long as they are made conditional on the hypothetical fact that the relevant terms denote something. To sum up, we use M^\forall , because it provides us with an ontologically neutral conditional syntax theory.

Let theory DT^* contain NFL, M^\forall and the Tarski bi-conditionals (1) but nothing else. Consider again scheme (2). Note that the following is an instance of axiom scheme **A7**:

$$T^\ulcorner(\phi \vee \neg\phi)^\urcorner \rightarrow E!^\ulcorner(\phi \vee \neg\phi)^\urcorner.$$

It tautologically follows from the above that $E!^\ulcorner(\phi \vee \neg\phi)^\urcorner$. It follows by **A4** and **A5** that

$$^\ulcorner(\phi \vee \neg\phi)^\urcorner = ^\ulcorner(\phi \vee \neg\phi)^\urcorner.$$

Hence, by the existential counterpart of axiom scheme **A4**, it follows that $\exists x(x = x)$. In other words, at least one thing exists. Theory DT^* is not ontologically conservative over NFL: just consider a model with an empty domain of quantification. Therefore, the theory DT^* is not deductively conservative over NFL. With scheme (1) *inclusive* NFL collapses into non-inclusive NFL. Note that axiom scheme **A7**, which is included in negative free logic but excluded in positive free logic, is essential for the argument.⁶

⁶An anonymous reviewer rightly pointed out that this argument can be generalized. Consider a theory S that extends NFL and M^\forall with the following rule: if S proves that ϕ , then S proves that $\rho^\ulcorner\phi^\urcorner$, where ρ is some special predicate. Then S proves that at least one thing exists. This is a nice generalisation, but that it will depend on the interpretation of the special predicate ρ whether the rule is sound and whether ontological non-conservativeness matters. As for the former, consider the predicate ‘is false’. As for the latter, consider the predicate ‘is a truth that can bloat the ontology of S ’.

We shall now prove that $E!t$, for all terms $t \in \mathcal{L}_A$, which are also all the terms of the language of DT^* . The proof is by induction on the complexity of terms. First, we need to prove that $E!0$. We have already shown that $E!\ulcorner(\phi \vee \neg\phi)\urcorner$, so one can use axiom scheme **A4** to infer from **M5** that:

$$\ulcorner(\phi \vee \neg\phi)\urcorner \cdot 0 = 0.$$

It follows from axiom scheme **A7** and the above that $E!0$. Second, we need to prove that $E!t'$, for all terms t . Assume that $E!t$. Use axiom scheme **A4** and $E!t$ to derive from **M4** that:

$$t+t' = (t+t)'$$

Successive application of axiom schemes **A7** and **A8** yields $E!t'$. Third, we have to prove that $E!(t+u)$, for all terms t, u . Assume that $E!t$ and $E!u$. Then derive from **M4** that:

$$t+u' = (t+u)'$$

It follows by axiom scheme **A7** that $E!(t+u)'$, whence it follows by axiom schemes **A5** and **A4** that $(t+u)' = (t+u)'$. Next, on the basis of **M2**, **A4** and $E!(t+u)'$ one can derive that:

$$(t+u)' = (t+u)' \rightarrow t+u = t+u.$$

From what we have already derived it follows that $t+u = t+u$, which by axiom scheme **A7** entails that $E!(t+u)$. Fourth and finally, we need to prove that $E!(t \cdot u)$, for all terms t, u . Assume once more that $E!t$ and $E!u$. Use the assumption and **M6** and **A4** to derive that:

$$t \cdot u' = (t \cdot u) + t.$$

By axiom schemes **A7** and **A8** it follows from the latter that $E!(t \cdot u)$. This concludes the proof.

The ramifications of the above result are threefold. First, DT^* proves the existence of many more beings than the base theory: it entails the existence of *infinitely* many entities rather than not even entailing the existence of a single entity. The reader can easily verify this on the basis of the above result and M^\forall -axioms **M1** and **M2**. Second, DT^* proves the existence not only of Gödel numbers that code sentences (e.g. $\ulcorner\phi \vee \neg\phi\urcorner$) but also of all *natural numbers*. Third, given the background of M^\forall and the Tarski-biconditionals (1), we obtain a full-blown collapse of NFL into CL. For what distinguishes the latter from (non-inclusive) free logic is **A4c**. Clearly, **A4** and $E!t$ yields **A4c**.

3 Philosophical reflections

By adding the Tarski biconditionals to inclusive negative free logic and the universal closure of minimal arithmetic, which is by itself an ontologically neutral combination, one can suddenly prove that at least one thing exists. This conclusion can be strengthened to the conclusion that infinitely many things exist and those are not all Gödel codes of sentences but rather all natural numbers. Against this background inclusive negative free logic collapses into non-inclusive free logic, which collapses into classical logic. At this point it seems that we face a choice:

- (1) We reject ontological deflationism on the basis of the fact that the Tarski biconditionals in combination with inclusive negative free logic and the universal closure of minimal arithmetic have ontological consequences;
- (2) We accept ontological deflationism and
 - (a) restrict the disquotational theory of truth, or
 - (b) move to other free logics, or
 - (c) weaken arithmetic.

Let us discuss the options.

The first option is to reject ontological deflationism on the basis of the fact that the Tarski biconditionals in combination with inclusive negative free logic and the universal closure of minimal arithmetic have ontological consequences. We have seen that the universal closure of minimal arithmetic is not needed to derive that at least one thing exists, which cannot be derived in inclusive logic only. Yet, to paraphrase Tarski, what is relevant is that we don't have to give any ontological attitude *we may have had*. Nobody denies that at least one thing exists, so it is reasonable to claim that the spirit of ontological deflationism has not been violated yet. However, there is more at stake. The claim that natural numbers exist is incompatible with various forms of mathematical anti-realism and the claim that infinitely many natural numbers exist is incompatible with finitism. Here we do have clear cases of a violation of the spirit of ontological deflationism.

One way out (2a) is to restrict the disquotational scheme as follows:

$$E!^{\ulcorner} \phi^{\urcorner} \rightarrow (T^{\ulcorner} \phi^{\urcorner} \leftrightarrow \phi). \quad (6)$$

This blocks the argument for DT* not being ontologically conservative. Moreover, it is a move that is akin to the ontological neutralisation of minimal arithmetic: scheme (6) is by **A4** a consequence of the vacuous universal closure of scheme (1), i.e. $\forall x (T^{\ulcorner} \phi^{\urcorner} \leftrightarrow \phi)$ for *sentences* ϕ . Let us see whether this is a good idea.

Suppose that ϕ and $\neg E!^{\ulcorner} \phi^{\urcorner}$. (If not, scheme (6) collapses into scheme (1), since it is still provable that, if $T^{\ulcorner} \phi^{\urcorner}$, then ϕ .) In that case one has by **A8** ϕ and $\neg T^{\ulcorner} \phi^{\urcorner}$, which seems odd.⁷ Worse still, it creates a tension between the *inner logic* and the *outer logic*.

Worse still, it creates a tension between the *inner logic* and the *outer logic*. Let theory DT** contain NFL, M^{\forall} and the restricted Tarski-biconditionals (6). The outer logic is the set of sentences that can be proved in DT**. It contains the law of excluded middle. The inner logic of DT** is the set of sentences such that it is provable within the system that they are true ($IDT^{**} =_{df} \{\phi \mid DT^{**} \vdash T^{\ulcorner} \phi^{\urcorner}\}$). We will now show that it does not contain the law of excluded middle. In order to do so we need to tap into the expressive resources of minimal arithmetic. In particular, we will need the Gödelian disjunction function *disj*, which takes the Gödel codes of two sentences to the Gödel code for the disjunction of those sentences – see (Boolos et al.,

⁷Admittedly, it is less of a problem than a theory of truth that *proves* $\phi \wedge \neg T^{\ulcorner} \phi^{\urcorner}$, for some sentence ϕ . See (Horsten, 2011, p. 127). Here we have only that it is *satisfiable* that ϕ and $\neg T^{\ulcorner} \phi^{\urcorner}$.

2003, Prop. 15.1). Let us use a primitive function symbol $\dot{\vee}$ to denote the Gödelian disjunction function – see (Feferman, 1966) for the idea of dot notation. Then we have that:

$$\ulcorner \phi \dot{\vee} \psi \urcorner = \ulcorner \phi \vee \psi \urcorner. \quad (7)$$

An immediate consequence of scheme (7) and **A6** is that:

$$T \ulcorner \phi \dot{\vee} \psi \urcorner \leftrightarrow T \ulcorner \phi \vee \psi \urcorner. \quad (8)$$

Adding scheme (8) to DT^{**} is ontologically conservative. In a nutshell, the argument starts from the assumption that $\neg E! \ulcorner \phi \urcorner$, subsequently derives that $\neg E! \ulcorner \phi \dot{\vee} \neg \phi \urcorner$, continues with deducing that $\neg T \ulcorner \phi \dot{\vee} \neg \phi \urcorner$ and, hence, $\neg T \ulcorner \phi \vee \neg \phi \urcorner$. Therefore, the inner logic does not contain the law of excluded middle. It looks as if the strategy of restricting the Tarski-biconditionals can only work if one also restricts the expressive resources of languages extending the language of arithmetic, viz. by disallowing a primitive function symbol that denotes the Gödelian disjunction function.

Another way out (2b) consists in moving to other free logics. This can be achieved by a restriction of axiom scheme **A7**, so that instances with the truth predicate are not allowed. Subject-predicate sentences without the truth predicate still imply that the terms they contain do refer, but subject-predicate sentences with the truth predicate no longer imply that names of sentences refer. This restriction blocks the argument for DT^* not being ontologically conservative, but it also moves us into the territory of *positive* free logic, which allows for certain subject-predicate sentences to be *true* even if some of the terms that they contain do not refer. Of course, we can also move straight to positive free logic by dropping **A7** and **A8** altogether. In any case, this option comes down to a rejection of *semantical* deflationism: the adherence to the disquotational theory of truth forces one to reconsider the semantics of subject-predicate sentences with non-denoting terms. This may be embraced by some philosophers, since truth itself may be seen as a semantical notion. Still, it is remarkable how strong the effect of the disquotational theory of truth is.

A final way out (2c) is to weaken arithmetic to such an extent that the ontological consequences that go against mathematical anti-realism or finitism can be avoided. One instantiation of this strategy that comes to mind is the so-called ‘purely modal strategy’ employed by nominalists and discussed in (Burgess and Rosen, 1997). In this context one can employ a purely modal strategy by adding the possibility and necessity operators to the language, adding a background modal logic and prefixing the axioms of $M^{\dot{\vee}}$ with possibility operators. Even considering that one can easily derive that necessarily at least one thing exists, not the actual but merely the possible existence of the natural numbers will be derivable. However, there is a trade-off: while one can avoid the ontological consequence, the usefulness of the weaker theory of arithmetic as a theory of syntax will be doubtful. Indeed, with the weaker theory of arithmetic in the background it is not clear whether the Tarski biconditionals are about *truth at the actual world*. Similarly, suppose that one jumps on the possibility to have partial function terms in free logic (Nolt, 2014, Section 5.2) to turn the successor function into a partial function, i.e. for some (large) natural number no successor is defined. But surely that will affect the theory’s ability to represent syntax.

Acknowledgements. We would like to thank the audiences at the Second CLAW-DWMC Symposium (22 April 2016, Leuven) and FilMat 2016 (26-28 May 2016, Chieti) for their comments on an early version of this paper. We would also like to thank the anonymous reviewers for their insightful comments.

References

- Boolos, G. S., J. P. Burgess, and R. C. Jeffrey (2003). *Computability and Logic* (Fourth ed.). New York: Cambridge University Press.
- Bricker, P. (2014). Ontological commitment. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 ed.).
- Burgess, J. P. and G. A. Rosen (1997). *A Subject with No Object: Strategies for Nominalistic Interpretation of Mathematics*. Oxford University Press.
- Feferman, S. (1966). Arithmetization of metamathematics in a general setting. *Journal of Symbolic Logic* 31(2), 269–270.
- Field, H. (1999). Deflating the conservativeness argument. *Journal of Philosophy* 96(10), 533–540.
- Halbach, V. (2001). How innocent is deflationism? *Synthese* 126(1-2), 167–194.
- Horsten, L. (1995). The Semantical Paradoxes, the Neutrality of Truth and the Neutrality of the Minimalist Theory of Truth. In P. Cortois (Ed.), *The Many Problems of Realism (Studies in the General Philosophy of Science: Volume 3)*. Tilberg University Press.
- Horsten, L. (2011). *The Tarskian Turn. Deflationism and Axiomatic Truth*. MIT Press.
- Horwich, P. (1998). *Truth*. Clarendon Press.
- Ketland, J. (1999). Deflationism and Tarski's Paradise. *Mind* 108(429), 69–94.
- Nolt, J. (2014). Free logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2014 ed.).
- Shapiro, S. (1998). Proof and Truth: Through Thick and Thin. *Journal of Philosophy* 95(10), 493–521.
- Tarski, A. (1936). The concept of truth in formalized languages. In A. Tarski (Ed.), *Logic, Semantics, Metamathematics*, pp. 152–278. Oxford University Press.
- Tarski, A. (1944). The semantic conception of truth: And the foundations of semantics. *Philosophy and Phenomenological Research* 4(3), 341–376.