Journal of Applied Non-Classical Logics
Vol. 00, No. 00, Month 201X, 1–16

Truth without Standard Models: some conceptual problems reloaded.

(*Received 00 Month 201X; final version received 00 Month 201X*)

A theory of truth is usually demanded to be consistent, but ω-consistency is less frequently requested. Recently, Yatabe (Yatabe, 2011b) has argued in favor of ω-inconsistent first-order theories of truth, minimizing their odd consequences. In view of this fact, in this paper we present five arguments against ω-inconsistent theories of truth. In order to bring out this point, we will focus on two very well-known ω-inconsistent theories of truth: the classical theory of symmetric truth FS and the non-classical theory of naïve truth based on Łukasiewicz infinitely-valued logic: PAŁT.

Keywords: theories of truth; non-standard models; ω-inconsistency

1. Introduction

In this paper we argue against ω-inconsistent theories of truth. Although Leitgeb (Leitgeb, 2001), Field (Field, 2008) and [anonymized] (anonymized, 2010) have criticized ω-inconsistent theories of truth, this issue continues to be controversial. In fact, for instance, Yatabe (Yatabe, 2011b)(Yatabe, 2015), among others, has recently argued in favor of ω-inconsistent first-order theories of truth. In his papers, he focuses on the technical side of the matter and argues that rejecting ω-inconsistent theories of truth would involve to reject co-induction, which is a mathematical principle (dual to induction) useful to prove properties of infinite structures, e.g. infinite streams, infinite trees, infinite process or infinite data structures[1]. Because of that, we think it is important to put under consideration some undesirable philosophical features of ω-inconsistent theories of truth in order to emphasise the adverse consequences of accepting such theories[2]. Then, we are going to offer some new reasons to give up Yatabe's point of view. For doing so, we will focus on very well-known theories of this kind: the classical theory of symmetric truth FS and the non-classical theory of naïve truth based on Łukasiewicz infinitely-valued logic: PAŁT. We present both systems over first-order arithmetic[3] and identify five conceptual problems concerning ω-inconsistency[4].

The rest of the article is organized as follows: in the next section we introduce some technical preliminaries and then review general features of FS and PAŁT. In section 3, we provide five reasons connected to ω-inconsistency against their acceptance. We argue that a theory of truth for arithmetic which is ω-inconsistent cannot capture the

[1]Roughly, and connected with the concept of truth, Yatabe shows how to identify paradoxical sentences with objects co-inductively defined. In section 3, we will analyse in detail this relation.
[2]In general the reasons for discarding ω-inconsistent theories of truth have centered on the alterations on the ontology of the interpretations of the base theories they produce ((Leitgeb, 2001), (Field, 2008) and [anonymized] (anonymized, 2010)). Of course we will consider this kind of arguments, and also we will provide additional considerations in Section 3.1.
[3]We will consider theories of truth over arithmetic, but applications to other more comprehensive base theories are presumably intended for these theories. Arithmetic is a convenient setting, since by fixing some Gödel coding it can express its own syntax, which is needed for the formulation of a truth system.
[4]Although we will focus on this two particular theories, as suggested by an anonymous referee , some of the problems presented here may be directly extended to other consistent but ω-inconsistent theories of truth.

intended ontology of the base theory and, therefore, it is not able to provide a better understanding of the intended model. Further, in this kind of systems the extension of the truth predicate will always contain objects that are not (codes of) sentences, among others. Also, as mentioned, we choose arithmetic in order to provide a theory of syntax, but, as in the case of the truth predicate, ω-inconsistency entails that syntactic predicates, such as $Sent$, will always contain objects that are not codes of sentences. And, finally, ω-inconsistent theories such as FS and PAŁT have important counterintuitive consequences regarding sentences containing the truth predicate. This casts doubts on whether the predicate in question is, after all, a truth predicate for that language[5].

2. Truth for first-order arithmetic

2.1 *Technical preliminaries*

Let $\mathcal{L}$ be the usual first-order language of arithmetic, with 0 as its only individual constant, S as a monadic function symbol for the successor function, $+$ and $\times$, and finitely many function symbols for primitive recursive functions. Let $\mathcal{L}_T$ be $\mathcal{L}$ plus the new monadic predicate symbol T. PAT is the first-order Peano arithmetic, PA, formulated in $\mathcal{L}_T$ with all the instances of induction given by $\mathcal{L}_T$-formulae. Let N be $\mathcal{L}$'s intended model, and let ω be its domain. By 'standard' or 'intended model of $\mathcal{L}_T$' we mean any model whose restriction to $\mathcal{L}$ is N.

For each $n \in \omega$, the term consisting of n occurrences of S followed by 0 is its canonical name or numeral in $\mathcal{L}_T$. We note it $\bar{n}$, as usual. We assume a fixed Gödel coding of every string of symbols σ of the vocabulary of $\mathcal{L}_T$ with a number $\#(\sigma) \in \omega$. We write $\ulcorner \sigma \urcorner$ for the numeral of the code of σ. We use the usual dot notation to denote certain primitive recursive syntactic functions: if A is a sentence of $\mathcal{L}_T$, $\underset{\cdot}{\neg}$ denotes the mapping that sends the code of A to the code of $\neg A$. $\underset{\cdot}{\wedge}$ and $\underset{\cdot}{\forall} v$ denote analogous operations for conjunction and the universal quantifier. $\dot{x}$ denotes the function that maps any number n to the code of its numeral. Let $x(y/z)$ denote the substitution function, that applied to the code x of a formula A and the codes y and z of terms t_1 and t_2, respectively, returns the code of the formula that obtains by replacing t_2 in A with t_1. As usual, we write $\ulcorner A(\dot{x}) \urcorner$ as short for $\ulcorner A(x) \urcorner(\dot{x}/\ulcorner x \urcorner)$ to bind x from outside corner quotes.

The set of atomic sentences of $\mathcal{L}_T$ is recursive, as well as the set of true atomic sentences, and also the sets of sentences and variables of $\mathcal{L}_T$. Thus, these sets will be represented in the theory by the following formulae: $At(x)$, $Ver(x)$, $Sent(x)$ and $Var(x)$, respectively.

2.2 *FS*

FS is an axiomatic system formulated in a Hilbert-style calculus. It was introduced by Friedman and Sheard (Friedman & Sheard, 1987), and thoroughly studied by Halbach

[5]It's worth noting that this relation between ω-inconsistency and non-standard models holds only over consistent theories. For instance, an arithmetical theory of truth built over the paraconsistent logic (LP) developed by Graham Priest, although is ω-inconsistent, does not disturb the standard ontology of PA. In other words, in the models, the intersection between the extension of the truth predicate and the extension of the negation of the truth predicate can be non-empty, although the restriction to the arithmetical vocabulary can be the standard model. Another more recently example is Fjellstad (Fjellstad, 2016), who has shown that the non-transitive theory of truth STTT, developed in (Cobreros, Égré, Ripley, & Van Rooij, 2014) is ω-inconsistent, although doesn't rule out the standard model. This theory is also inconsistent, because it proves $\Rightarrow \lambda$ and $\Rightarrow \neg\lambda$, for some sentence λ (e.g. this is the case of the liar sentence), although it's not paraconsistent, because explosion is a valid inference. In view of these facts, in what follows we'll focus only on consistent theories of truth.

(Halbach, 1994), (Halbach, 2011)[1]. It represents an approach to truth, by treating it as a primitive predicate, governed by certain axioms and rules.

FS is a type-free (or self-referential) theory of truth, *i.e.*, its truth predicate is intended to apply not just to sentences from the base theory, but also to sentences containing the truth predicate. On the one hand, this entails much more expressive power than typed theories, getting one closer to the behavior to the truth predicate in natural languages. On the other hand, this entails much more deductive power and to avoid a hierarchy of languages puts the theory at risk of inconsistency, due to Tarski's theorem on the undefinability of truth. Therefore, FS replaces the T-schema with two corresponding—weaker—rules: one that allows to prove the truth of any theorem of the theory—usually called 'rule of necessitation' or 'NEC' for short; and its converse, that allows one to remove the truth predication from a theorem—called 'rule of co-necessitation' or 'CONEC' for short.

Being closed under both rules entails that all provable sentences are provably true and, vice versa, all provably true statements are also provable themselves. Thus, in FS the internal and external logics coincide, *i.e.*, FS is a symmetric system[1]. In addition, it contains axioms that are the result of expressing the compositional clauses in Tarski's definition of truth *in and for* the language of the theory. Therefore, FS is a compositional theory of *classical* truth, since according to Tarski's clauses, for every sentence, either it or its negation must be true, but not both[2].

Once the base system is chosen, it becomes clearer how to formulate the axioms and rules of FS, as they depend on the syntax of the language of the base theory. We choose one possibility among many: first-order Peano arithmetic.

Let FS be the theory consisting of PAT plus the following axioms[3]:

$(AT)\ \forall x(At(x) \to (Tx \leftrightarrow Ver(x)))$[4]
$(T\neg)\ \forall x(Sent(x) \to (T \underset{\cdot}{\neg}\, x \leftrightarrow \neg Tx))$
$(T\wedge)\ \forall x, y(Sent(x) \wedge Sent(y)) \to (T(x \underset{\cdot}{\wedge} y) \leftrightarrow (Tx \wedge Ty)))$
$(T\to)\ \forall x, y(Sent(x) \wedge Sent(y) \to (T(x \underset{\cdot}{\to} y) \leftrightarrow (Tx \to Ty)))$
$(T\forall x)\ \forall x, v(Sent(\underset{\cdot}{\forall} vx) \to (\forall t T(x(t/v)) \leftrightarrow (T(\underset{\cdot}{\forall} vx))$

plus the inference rules:

[1]Friedman and Sheard's original axiomatization is slightly different from Halbach's formulation, which is going to be utilized here. Also, they did not call their system 'FS'. However, both theories are logically equivalent (see (Halbach, 1994) for details).

[1]According to Leitgeb, (Leitgeb, 2001, p. 282), when one refers to the outer and the inner logic of a theory of truth, what one means is that the logical laws in such theories can show up in two different contexts: outside of applications of the truth predicate and inside of such contexts. For example, Leitgeb says:

> "[...]there are consistent theories of truth in which both sentences of the form A or not A and not Tr(A or not A) are derivable. While the former is an instance of the classical law of the excluded middle, the latter denies an instance of the excluded middle in the context of the truth predicate. Accordingly, although the outer logic of the theory might be genuinely classical, its inner logic certainly is not. This is in contrast with Tarskis theory, which is an example of a theory of truth for which the outer and the inner logic coincide (they are both classical).

[2]As an anonymous referee suggested, the second part of the sentence does not suffice to exclude substructural theories (which presumably count as non-classical). However, for the sake of simplicity here we will not consider this kind of theories.

[3]We take the following axioms for T from (Halbach, 1994).

[4]Let's recall that Ver(x) is a predicate that means that x is (the code of) a true atomic sentence.

$$(\text{NEC}) \quad \frac{\vdash A}{\vdash T\ulcorner A \urcorner} \qquad\qquad (\text{CONEC}) \quad \frac{\vdash T\ulcorner A \urcorner}{\vdash A}$$

As mentioned, FS is a self-referential, compositional and symmetric theory of classical truth. It implies PA's consistency statement and, thus, its Gödel sentences and a global reflection principle for PA. It is important to mention that FS is a subtheory of the revision theory of nearly stable truth $T^{\#}$ with N as its base model (for details, see Gupta and Belnap (Gupta & Belnap, 1993)).

Halbach (Halbach, 1994) proves the consistency and, thus, satisfiability of FS, by showing that ω-models for subsystems of FS with a limited number of applications of NEC and CONEC can be obtained using revision semantics. As a consequence, FS is arithmetically sound: it does not prove any arithmetically false statement. All of these features make FS an attractive theory of truth.

2.3 *PAŁT*

PAŁT was first introduced by Restall (Restall, 1992) (and, in a different way, by Hàjek et al. (Hàjek, Paris, & Shepherdson, 2000)). The base logic of this theory, unlike FS, is Łukasiewicz predicate logic: ∀Ł. The most remarkable features of this logic, compared with classical logic and another multivalued logics, are related to its conditional: the logic doesn't validate the propositional law of contraction, but validates *modus ponens*. Restall (Restall, 1992) and Hàjek et al. (Hàjek et al., 2000) have shown that, among others, these characteristics allow us to build a consistent naïve truth theory—i.e. with the full T-schema. In fact, not only we can assign a stable semantic value to the liar sentence, but also we can have a suitable conditional, avoiding the Curry paradox (see Bacon (Bacon, 2013) for details).

First, we will formally present the logic ∀Ł, and then introduce the theory PAŁT.

2.3.1 Łukasiewicz predicate logic: ∀Ł

We introduce Łukasiewicz's predicate logic in a semantic fashion, as usual (cf. Hàjek (Hàjek, 1998), Yatabe (Yatabe, 2011b))[1]. The set of semantic values is the real interval [0,1] and 1 is the only designated one. Let $M = \langle D, r_P, m_c \rangle$, where $D \neq \emptyset$, $r_P : D^n \to [0,1]$ (where 'n is the arity of the predicate) and $m_c \in D$. For each valuation v from the object variables into the domain D and every two formulae A and B of the language we define $\|A\|_{M,v}$ to be the truth value of A in M, v:

- $\|P(x, ...c, ...)\|_{M,v} = r_P((v(x), ..., m_c, ...))$, where P is a predicate, c is a constant and x is a variable.
- bottom: $\|\bot\|_{M,v} = 0$
- negation: $\|\neg A\|_{M,v} = 1 - \|A\|_{M,v}$
- fusion: $\|A \circ B\|_{M,v} = max\{\|A\|_{M,v} + \|B\|_{M,v} - 1; 0\}$
- conjunction: $\|A \wedge B\|_{M,v} = min\{\|A\|_{M,v}; \|B\|_{M,v}\}$
- existential: $\|\exists x A(x)\|_{M,v} = supr\{\|A(x)\|_{M,v'} : v(y) = v'(y), \textit{for all variables y except possibly } x\}$[2]

Other connectives, like disjunction ∨ and fission ↑ can be defined in a natural way: $\|A \vee B\|_{M,v} = max\{\|A\|_{M,v}; \|B\|_{M,v}\}$ and $\|A \uparrow B\|_{M,v} = min\{\|A\|_{M,v} + \|B\|_{M,v}; 1\}$. Perhaps the most interesting is the case of the conditional. It's defined as the residuum of the fusion connective:

[1]∀Ł is not recursively axiomatizable. See, Hàjek (Hàjek, 1998) for details.
[2]Where v' is a valuation that is equal to the valuation v, but can differ in the assignation of x.

- $\|A \to B\|_{M,v} = min\{1; 1 + \|A\|_{M,v} - \|B\|_{M,v}\}$

It's trivial to check that the conditional takes the designated value if and only if the value of the antecedent is less than or equal to the value of the consequent. Also, it's worth noting that we can find counterexamples for the propositional law of contraction, and *reductio*. Nevertheless, as we mentioned, the logic validates *modus ponens* and a kind of theorem of deduction (for details, see Hàjek (Hàjek, 1998)).

2.3.2 The theory of truth: PAŁT

In order to build an arithmetical theory over this logic, we need to modify the axioms of PA. Thus, we use the axioms traditionally given in the context of classical logic, but we need to replace the axiom schema of induction by a rule (because the logic doesn't validate the propositional law of contraction):

$$A(\bar{0}), \forall x(A(x) \to A(x+1))) \vdash \forall x A(x)$$

The resulting theory is known as $PAŁ$. However, the set of theorems of PA (with the axiom schema) over classical logic and the set of theorems of $PAŁ$, replacing the axiom schema of induction by the induction rule, over Łukasiewicz logic are exactly the same (cf. Restall (Restall, 1992), Hàjek (Hàjek et al., 2000)). Therefore, for simplicity, we will refer to both theories as PA[1].

The next step is to augment the language with an unary predicate T that satisfies the T-schema for all the sentences A of $\mathcal{L}_T$: $T(\ulcorner A \urcorner) \equiv A$. We call this theory PAŁT. It's worth noting that, due to the features of the underlying logic, the resulting theory is consistent. In other words, PAŁT is a naïve theory of truth that contains all instances of T-schema and it's not trivial[2].

3. Arguments against ω-inconsistent Theories of Truth

Though consistent, FS and PAŁT are ω-inconsistent[3]. Since both proofs of ω-inconsistency are widely known, we leave these to an appendix.

So, in what follows we will argue against ω-inconsistent theories of truth such as FS and PAŁT showing that ω-inconsistency has significantly negative philosophical consequences and therefore should be avoided.

Usually, two main concerns guide philosophical investigations by giving a theory of truth for some base theory or a model. On the one hand, it is highly desirable for a theory of truth to entail those and only those truth principles that capture the correct applications of the truth predicate. Tarski (Tarski, 1944) argued that any suitable theory of truth should imply every instance of the T-schema. However, a weaker criterion is to consider capture and release (NEC and CONEC), see, among others, (Beall, 2009).

On the other hand, due to Tarski's (Tarski, 1935) undefinability result, it is not possible to consistently embrace the unrestricted T-schema within a language allowing unrestricted self-reference in a classical frame.

[1]Restall (Restall, 1992) shows that with the standard interpretation, as a semantic consequence of the theory the models are bivalent. On the other hand, Hàjek et al. (Hàjek et al., 2000) add an axiom to the theory for identity. In both cases, we have that if $A(x)$ is an arithmetical formula, then $PAL \vdash A(x) \vee \neg A(x)$. Let's recall that given a theory T, we say that M is a model of the theory iff $\|A\|_M$=1 for each A, theorem of T.

[2]For a proof of non-triviality, see Hàjek (Hàjek, 1998).

[3]In this context, since we are working in an arithmetic frame, a system $\mathcal{T}$ is said to be ω-inconsistent if, for some formula A with exactly x free, both $\mathcal{T} \vdash \exists x \neg A$ and also $\mathcal{T} \vdash A(\bar{n})$ for each $n \in \omega$. As an anonymous referee suggested, it would not be a it would not be a sufficiently general definition if we were not working with arithmetic theories if we were not working with arithmetic theories.

In a non-classical frame, even though Tarski's Theorem doesn't necessarily apply, it's not just the T-schema what truth theorists want (see (Leitgeb, 2007) for some desiderata for a theory of truth). Classical and non-classical truth theorists also seek non-trivial systems avoiding unsoundness. However, we will argue that there is a third concern that should be taken into account: a good theory of truth must not exclude its intended base model, the intended interpretation of its base system.

While theories as FS and PAŁT consistently imply a good amount of intuitive truth principles as the compositionality of truth (FS), and the T-schema (PAŁT)[1] they fail to comply with this third requirement: their ω-inconsistency directly entails a deviation from N, PA's intended interpretation. Also, as we will argue, ω-inconsistency can often lead to unsoundness. From our point of view, this result has significant consequences regarding the capacity of FS to express truth. Nevertheless, this thought does not seem to have hindered the acceptance of those systems among truth theorists. For instance, Sheard (Sheard, 2001, p. 179) argues that:

> [...] [the] fascinating discovery that some consistent axiomatic theories of truth are in fact ω-inconsistent does not present a significant impediment to the effective use of those theories.

Recently, Yatabe (Yatabe, 2011b) shows that there is an interesting relation between ω-inconsistent theories of truth and co-induction. In his own words:

> ω-inconsistent truth theories are intrinsically equipped the machinery, co-induction, which is widely used in computer science and enables to represent infinite processes. Rejecting them involves rejecting co-induction at a time though it is very useful and natural, therefore we need to balance the profits and losses of rejecting co-induction before we reject them (Yatabe (Yatabe, 2011b, p.102))[2].

Let's clarify this quote comparing a co-inductive definition and an inductive one[3].

Example 1 (Inductive definition)**.** *Let K be any set. Then the list of K is of the type* $\langle K^{<\omega}, \gamma : (1 + K \times K^{<\omega} \to K^{<\omega})\rangle$*, such that:*

- ***Initial step****: 1 consists in the empty set* $\langle\rangle$
- ***Succesor step****: given* $k_0 \in K$ *and* $\langle k_1, ..., k_n\rangle \in K^{<\omega}$*,* $\gamma(k_0, \langle k_1, ..., k_n\rangle) = \langle k_0 k_1, ..., k_n\rangle$

In this sense, this inductive definition generates the possible lists of elements of K. On the other hand, co-inductive definitions allows to build infinite objects.

Example 2 (Co-inductive definition)**.** *Let A be any set. Then the streams of K is of the type* $\langle K^{\infty}, \gamma : (K^{\infty} \to K \times K^{\infty})\rangle$*, such that for any* $(\langle k_0, k_1 ...\rangle \in K^{\infty}\ \gamma(\langle k_0, k_1 ...\rangle) = (k_0, \langle k_1 ...\rangle) \in K \times K^{\infty}$*.*

The intuitive idea behind this definition is that the function takes the first element of the infinite stream and delays the construction of the object one step further. Induction corresponds to well-founded structures that start from a basis which serves as the foundation: e.g., natural numbers are inductively defined via the base element zero and the successor function. On the other hand, co-induction eliminates the initiality condition and keeps the successor step. Hence, the characteristic features of co-inductive construc-

[1] Note that if we add the T-schema to FS the theory becomes trivial. The same happens if we add compositionality to PAŁT.

[2] Yatabe seems to conclude that since there is a relation between ω-inconsistency and co-induction, if we reject a ω-inconsistent theory of truth, we are rejecting co-induction *tout court*. As an anonymous referee pointed out, the fact that co-induction is widely used in computer science is independent of accepting it in the context of theories of truth. Later we will argue against Yatabe's position.

[3] We take these two examples from (Yatabe, 2011b).

tions are (i) one only uses finitely many already-constructed objects to construct a new object, and (ii) one needs infinitely many steps to reach the initial construction case (this is not inductive construction). Then, there is no base case in a co-inductive process, and while it may appear circular, the structure is well formed since co-induction corresponds to the greatest fixed point interpretation of recursive definitions: namely, the set of all infinite lists of numbers. So, co-inductive structures are essentially infinite-length structures. It is well-known that co-inductive constructions play a very important role in computer science to represent behaviors of non-terminate automatons.

In this vein, Yatabe uses this technical machinery in order to show that the objects that represent the paradoxical sentences can be co-inductively defined as infinite streams[1].

Therefore, these views endure in the literature. In what follows, we will develop five arguments against ω-inconsistency that, contrary to what quotes above express, will show that ω-inconsistency should be avoided[2].

3.1 *The Semantical Argument*

From our point of view, any interesting theory of truth must be supported by philosophical reasons. Otherwise, it might not express a legitimate truth predicate[3]. One of the reasons is that adding a truth predicate to some base theory or model should not interfere with the intended ontology of that base theory or with that model, respectively[4]. In order to emphasise this, we recall Gupta and Belnap's (Gupta & Belnap, 1993, p. 142) slogan:

> The addition of a truth predicate to a language does not disturb the logical structure of the language in any way.

In theories as FS or PAŁT, if we are willing to talk about truth in a base theory, we think that we must have in mind some interpretation for the language of the theory: the intended interpretation. In fact, it would not make much sense to speak about the truth or falsity of uninterpreted formulae. Therefore, if we allow our truth principles to interfere precisely with the intended interpretation in a way that the resulting theory is not true in it, our theory of truth does not seem to be serving its original purpose.

According to us, an ω-inconsistent theory of truth is not capable of fulfilling this requirement. FS and PAŁT are special cases of this phenomenon: they disturb the intended arithmetical ontology, since they lack the standard model N. This means that their so-called truth predicates do not seem to express legitimate truth for their underlying theory, respectively, which was what we wanted in the first place. Let us emphasise the point we are concerned about. Of course, our worry is not the existence of non-standard models for FS or PAŁT. Our worry is that these type of theories only have non-standard models as a consequence of introducing axioms and rules that try to capture the concept of truth. Thus, ω-inconsistent theories of truth cannot conserve the

[1]For the technical details, see (Yatabe, 2011b).
[2]It has been shown that results for the second-order case are even worse. See (Picollo, 2013) or (anonymized, 2013).
[3]We agree with Hannes Leitgeb on this point: any successful formal theory of truth must be supported by philosophical argumentation. Nevertheless, some authors could object that, contrary to what we affirm, several truth-theories appeal to the uniformity of their truth-theoretical principles alone. Against this point, we reply that it seems complicated to accept formal principles without a philosophical evaluation at all. Obviously it is not clear how to divide the reasons to accept a theory in merely philosophical or a merely formal. In any case, we would like to discuss whether the philosophical ideas we are offering are good reasons for rejecting ω-inconsistent theories of truth. For details see (Leitgeb, 2001).
[4]A similar argument has been stressed by [anonymized] (anonymized, 2010) in the context of the discussion about Yablo's Paradox.

standard semantic values for at least some interpretation.

However, it could perhaps be argued that these ontological difficulties are not so relevant. For example, according to Yatabe (Yatabe, 2011b, p. 102)

> This analysis on ω-consistency [that implies that ω-inconsistency is caused by the fact that the truth predicate enables us to define formulae co-inductively], together with the fact that adding the truth predicate makes the theory's proof-theoretic strength very high, reconfirms the fact that the truth predicate increases the expression power of language. That is based on the very nature of truth predicate, and non-standardness is its direct consequence (the ontology is secondary problem in this sense).

It seems that Yatabe doesn't consider but just in this parenthetical note the ontological problem. Nonetheless, we would like to emphasize that we do not agree that interference is a secondary problem. The fact that the introduction of the concept of truth alters the nature of the models has important technical and conceptual implications. According to $\mathcal{L}$'s standard interpretation, relation symbols refer to relations between natural numbers and function symbols to functions defined over them. Quantifiers also have the natural numbers as their scope. In this way, the intended interpretation allows $\mathcal{L}$-expressions such as $\neg\exists x(Sx = 0)$ to 'say' something just about natural numbers. Of course, there are no completeness results for PA avoiding the existence of non-standard models for this system, but soundness is crucial: PA-axioms and theorems must be true in those structures whose domains are exclusively constituted by natural numbers, for otherwise they would no longer be capturing the intended ontology.

One of the goals when introducing a truth predicate to $\mathcal{L}$ is that $\neg\exists x(Sx = 0)$ comes out true precisely *because* it can be interpreted as saying something true about natural numbers . The fact that ω-inconsistent theories of truth lack the standard model means that, although sentences such as $T\ulcorner\neg\exists x(Sx = 0)\urcorner$ may be entailed by the theory, they can never be seen as expressing the truth of a sentence that concerns just natural numbers—of an arithmetically true statement—since sentences from the theory can be no longer seen as saying something true about ω. In this sense, an ω-inconsistent theory of truth defines truth conditions for formulae not depending (only) on the natural numbers. Thus, for us, its alleged truth predicate is not a legitimate truth predicate for arithmetic.

It could also be replicated that ω-inconsistency is an inevitable consequence of the expressive power of the concept of truth. In this line, as we mentioned, Yatabe (Yatabe, 2011b) indicates an interesting and promising relation between ω-inconsistency and co-induction. In his paper he analyzes what happens when a semantically closed language describes itself. According to him, co-induction is the machinery that allows to define the paradoxical sentences. The main claim of Yatabe is that ω-inconsistency is caused by the fact that the truth predicate allows to define formulae co-inductively. He insists that, generally speaking, co-inductive definitions involves non-standardness. In his paper, he defines a language in the metatheory whose domain is co-inductive formulae and then he shows that these co-inductive objects can be interpreted as the sentences that generate ω-inconsistency. Hence, these co-inductive (potentially infinite) objects defined by co-induction are infinite objects in the sense of the meta-theory, but finite (non-standard) objects in the sense of the object-theory.

As we will see later in our "Syntax Argument", in any non-standard model, there are sentences that are interpreted as infinite. In this sense, we can identify them with co-inductive formulae. Nevertheless, we would like to point out that the capability of identifying co-inductive formulae with sentences of a theory of truth *cannot be sufficient in order to cause* disturbance in the standard ontology of PA.

Firstly, there are several theories of transparent truth that allow co-inductive process and are conservative over the ontology of PA. The Kripke-Feferman theory of truth presented by Halbach (Halbach, 2011, pp. 200-201) is usually seen as an axiomatization of Kripke's semantic theory of truth and is conservative over the ontology of PA. Of course, no axiomatization of a kripkean fixed-point semantics can be complete in the sense of fully describing the standard model. But, as Halbach (Halbach, 2011, p. 217) shows, if the standard model is fixed as the underlying model of the language of the base theory, then KF fully characterizes the fixed points of Kripke's theory (even maximal fixed-points). In the same way, the paracomplete theory of truth presented by Hartry Field (Field, 2008) is conservative over the standard model of PA[1]. In other words, it is ω-consistent, and allows to identify some formulae with co-inductive formulae.

Secondly, let's see an example of a mock theory with a truth predicate that admits co-inductive definition of formulae but is ω-consistent. Let's add to PA the following sentence (a generalized version of the well known Truth-Teller): $A \leftrightarrow \forall x T\dot{h}(x, \ulcorner A \urcorner)$ where T is a truth predicate[1] and

$$\dot{h}(0, \ulcorner A \urcorner) = \ulcorner A \urcorner \text{ and } \dot{h}(S(\overline{n}), \ulcorner A \urcorner) = \dot{T}\dot{h}(\overline{n}, \ulcorner A \urcorner) \text{ for each } n \in \omega.$$

Intuitively, A is a consequence of transparency; in any theory for transparent truth, it would be desirable to have this equivalence for every sentence. Also, it's easy to define a co-inductive language and a co-inductive object whose intended meaning corresponds to this sentence[2].

However, it's trivial to note that there are extensions of the standard model of PA that are models of this theory[3].

In view of this, although it's very interesting the connection between co-induction and ω-inconsistency developed by Yatabe (Yatabe, 2011b), we claim that co-induction is not sufficient for ω-inconsistency. Therefore, there is not a direct connection between the concept of truth (and its expressive power) and non-stadardness. If it were the case, there would not be theories of truth keeping the standard ontology of the base theory and that accept co-inductive processes.

In sum, it does not seem a good idea to accept ω-inconsistent theories of truth. From a philosophical perspective, this type of theories produces deviations in the standard ontology of the base theory as a consequence of introducing the axioms that rule the behavior of the concept of truth. Moreover, if this alteration at the level of the nature of the models were a consequence of the complexity of truth, then all the theories capable of capturing this complexity would have to lack of standard models. But, as we have shown, this is not the case.

3.2 *The Better Understanding Argument*

As it's well known, Gödel's incompleteness theorems show that truth and provability in PA do not coincide, for there are many arithmetically true $\mathcal{L}$-formulae that are not provable in that system, such as the Gödel sentence for PA and its consistency statement[4]. Therefore, one might reasonably expect an axiomatic theory of truth for arithmetic to prove at least some of these arithmetically true $\mathcal{L}$-formulae to be true—though by Gödel

[1]Recently, Field (Field, 2015) combines the kripkean theory of truth with a variably strict conditional. There, he insists in imposing ω-consistency to the theory.
[1]Strictly speaking, add the relevant instances of the T-schema and the relevant sub-formulae of the sentence.
[2]Following the idea formally developed in (Yatabe, 2011b), it's possible to take $\dot{h}(\overline{n}, \ulcorner A \urcorner)$ as intuitively meaning $T\ulcorner ...T\ulcorner A\urcorner ...\urcorner$, with n iterations of T. In this vein, $\forall x T\dot{h}(x, \ulcorner A \urcorner)$ might be seen as an infinite iteration of the truth predicate, and so, as an infinite stream.
[3]For our goal, it was sufficient to add to PA this sentence. However, note that we could add an infinite number of sentences of this kind keeping standardness. Yatabe seems to confuse circularity and paradoxicality: A is circular and can be defined co-inductively, though it is not paradoxical.
[4]Here, as usual, we are assuming that PA is consistent.

incompleteness theorems, we know that no recursively axiomatizable consistent system would be able to prove them all[5].

It is usually thought that the more true-in-N statements a theory of truth for arithmetic entails, the closer it gets to truth for arithmetic—as long as it does not imply any falsity. For by proving arithmetically true formulae the system gets rid of many non-standard models (but not all of them) in which those formulae were false and, therefore, provides a better characterization of N[1]. Nonetheless, this is not always the case. There are truth theories for PA that imply many true-in-N statements PA does not and, notwithstanding, they do not provide a better understanding of PA's intended interpretation since they are not true in N.

FS is one of those theories. It is arithmetically sound and proves more true-in-N formulae than PA, including PA's Gödel sentence and its consistency statement. Also, it proves them to be true, by an application of NEC. In fact, we could add all truth-in-N sentences without getting into contradictions. So FS disposes of many non-standard models of PA. However, it does not provide a better characterization of N *because* it gets rid of this model too: although the more applications of NEC one allows the more arithmetically true formulae the theory proves, if arbitrarily many applications of the rule are allowed, one gets $T\dot{f}(\bar{n}, \ulcorner C \urcorner)$ for all $n \in \omega$. This leads to ω-inconsistency and, *a fortiori*, the lack of the standard model.

Thus, adopting FS as a truth theory implies giving up the possibility of gaining a better understanding of the underlying ontology of the base system. N does not belong to FS class of possible interpretations and, therefore, there is no sense in which we can think of FS as saying something about N, which is dismissed by the theory along with many non-standard interpretations. As a consequence, FS is completely incapable of expressing truth in this model—that is, truth over arithmetic—not even partially.

On the other hand, PAŁT is also arithmetically sound and proves more true-in-N formulae than PA. It can also be consistently extended with the set of arithmetic truths. However, as in the case of FS, the inability to capture the standard model implies that we can't interpret the theory as saying something about arithmetic. So, how could we get a better understanding of truth for arithmetic if we can't interpret the theory as saying something (true or false) about standard arithmetical operations between standard numbers?

Again, it might be questionable whether some assumptions in this argument are inconsistent with the adoption of a deflationary attitude about the notion of truth. Linked to this potential objection, we would like to emphasise that our point of view attempts to be neutral regarding the philosophical dispute between deflationary and robust conceptions about truth. That is, we would like the arguments we are giving have an impact on ω-inconsistent theories regardless of our point of view about the nature of truth. However, we would still like to stress the following. Nothing more distant to the de-

[5]As rightly a referee noted, it could be objected that it is not self-evidence that every axiomatic theory of truth for arithmetic has to prove at least some of the arithmetically true $\mathcal{L}$-formulae to be true. For example, some deflationist might disagree. Many philosophers believe that a deflationist theory of truth must conservatively extend any base theory to which it is added. This point of view is supported by Shapiro (Shapiro, 1998) and Ketland (Ketland, 1999). We agree that imposing a requirement of conservativeness could be a serious objection. But as Ketland also says: "Deflationism about truth is a pot-pourri" ((Ketland, 1999, p.69)]). And, in this case, it is crucial the transition from some thesis as 'truth is insubstantial' to 'truth theories must be conservative'. However, it is not clear to us that deflationism requires conservativeness, (for example, Field (Field, 1999) rejects such connection). In any case, for the current discussion our assumption does not seem problematic: the FSs supporter should share with us the fulfillment of this desiderata: it proves some $\mathcal{L}$-formulae, as Gödel sentence, that are true.

[1]As an anonymous referee suggested, one could object that there are uncountable many non-standard models and therefore one is not really better off in approximating N. Of course, this is right. It's not possible to eliminate all non-standard models. However, our point is that ω-inconsistent theories of truth as FS do not give a better understanding of what is arithmetically true *because* of the lack of the standard model.

flationist than to commit to the following thesis: the introduction of a truth predicate in PA causes a deviation in the intended ontology of the mentioned theory. If truth is not a robust property, as many deflationists hold, or if it is not a property at all, how can the addition of that notion to a base theory have the power to alter the class of the models in such a way that the resulting theory has only non-standard models? If the truth predicate really lacks all explanatory power, it should not have the power to eliminate the class of the intended models.

3.3 *The 'Extension of T' Argument*

A theory of truth should guarantee the application of the truth predicate just to designed truth bearers, such as propositions, sentences or, as in our case, numbers that codify those sentences. This requirement is better known as the item two of Tarski's (Tarski, 1935, pp. 187-188) Convention T.

Of course, no first-order recursively axiomatizable theory of truth for PA that allows infinitely many objects in the extension of the truth predicate is able to provide a truth predicate whose extension contains just codes of sentences in every model. However, it is reasonable to expect that the theory has at least one model (the standard one), such that the extension of the predicates is the one intended. If we adopt an ω-inconsistent theory of truth such as FS or PAŁT, the interpretation of the alleged truth predicate does not seem to be appropriate in this sense. In the case of FS, due to McGee's theorem (see the Appendix for details), we know that this system entails $T\dot{f}(\bar{n}, \ulcorner C \urcorner)$ for each $n \in \omega$ and, at the same time, $\neg\forall x T\dot{f}(x, \ulcorner C \urcorner)$. Hence, by $T\neg$, we must have a non-standard number c in the domain satisfying $T\dot{\neg}\dot{f}(x, \ulcorner C \urcorner)$. This means that there must an object c' in the extension of T that is the denotation of $f(c, \#(C))$. This object must be a non-standard number as well, by the injectivity of f.

In the case of PAŁT, also we show (see the Appendix) that there is a formula $T(\dot{g}(x, \ulcorner D \urcorner))$ that is not true for any standard number, but it's satisfied by a non-standard number, because $\exists x T(\dot{g}(x, \ulcorner D \urcorner))$ has the designated value. Hence, for any model, there must be a non-standard number c in the extension of T, such that $T(\dot{g}(c, \ulcorner D \urcorner))$. Therefore, in both systems we have the same consequence: there will always be a non-standard number in the extension of T. Yatabe (Yatabe, 2011a) acknowledges the problem:

> T interprets arithmetical operations on Gödel codes [...] to actual operations on formulae

but he doesn't extract the philosophical consequences: the truth predicate has non-standard numbers in its extension in every interpretation. In other words, coding is one-to-one and every $\mathcal{L}_T$-formula is codified by a standard number; so non-standard numbers cannot codify any $\mathcal{L}_T$-sentence. In other words, for every model, T's extension contains more than formulae, committing the theories to the existence of true objects in the models that are not intended to be truth bearers. This is a direct consequence of excluding the standard model among the possible interpretations, and therefore of lacking the intended intepretation. Thus, both systems seem to offer a characterization of a non-standard concept of truth.

3.4 *The Syntax Argument*

As mentioned above, theories of truth are usually based on PA because it's strong enough to represent the syntax of the theory. Once we add a truth predicate to a theory, we want to predicate truth to its true sentences, and thus we need to be able to speak about the sentences of the theory inside the theory. And PA allows for it. In other words, we can build predicates that represent syntactic concepts in the object language. For instance, we have used the predicate $Sent(x)$ to represent the sentences of $\mathcal{L}_T$.

In this line, the intended interpretation of $Sent(x)$ is the set of (codes of) sentences of the language. However, if we dispose the standard model this possibility is lost. In any non-standard model, if a formula with a free variable is satisfied by infinitely many object, then there must be a non-standard number in the domain of the model that satisfies the formula[1]. Therefore, any syntactic predicate such as $Sent(x), At(x)$, etc. is satisfied by a non-standard number at each non-standard model. Let us insist that the problem here is not that in any non-standard model the syntactic predicate $Sent(x)$ is satisfied by objects that don't represent sentences. The problem here is that, once we exclude the standard model, the extension of $Sent(x)$ will contain non-standard numbers in any model.

Many theorists have argued that, since these objects satisfy the predicate $Sent(x)$, though they are not sentences in the usual sense, they are non-standard sentences and, thus, still sentences. In this sense, one of the most important contributions to this topic is the famous Abraham Robinson's paper (Robinson, 1963)[2]. In his work, Robinson shows how to build theories capable to express their own syntax properties using non-standard models. Actually, in this kind of theories it's possible to express sentences of any length, including infinite sentences. However, it would commit us to non-standard infinitary languages, which are beyond the scope of the target theories presented on this paper. So, as they do not codify any $\mathcal{L}_T$-expression, we will not follow this route[3].

So far, we had showed that FS and PAŁT, due to their ω-inconsistency, cannot be interpreted as saying anything about the intended model of arithmetic, don't provide a better understanding of the standard model of arithmetic, and include non-standard elements in the extension of the truth predicate. Moreover, now, we have additionally shown that, as a consequence of ω-inconsistency, we can't interpret our theories as codifying adequately their own syntax.

3.5 *The Unsoundness Argument.*

A good theory of truth is one that entails as many intuitive truth principles as possible. But most importantly, it should not imply counterintuitive statements involving truth. Next we will show that FS and PAŁT do so as a consequence of being ω-inconsistent.

For the case of FS consider the following $\mathcal{L}_T$-sentence:

(RFL_{FS}) $\forall x(Bew_{FS}(x) \rightarrow Tx)$

where Bew_{FS} is the provability predicate for FS, expressible in this system. RFL_{FS} is a global reflection principle for FS: it states that all theorems of FS are true. This principle should be desirable to anyone embracing FS, for it establishes the soundness of FS. Moreover, it appears to be true according to FS since, by NEC, T applies to every theorem of FS.

However, FS proves the negation of RFL_{FS} and its falsity[4]. $\neg RFL_{FS}$ is a highly counterintuitive principle but—worst of all—it is strictly false. Although FS doesn't prove any arithmetically false statement, they entail incorrect truth-theoretical principles[5].

[1]See Kaye (Kaye, 1991) for a proof.
[2]We are very grateful to an anonymous referee who pointed out this paper and its relevance in this subject.
[3]Field in (Field, 2015) also argues against non-standard models based on syntactic interference.
[4]See Halbach and Horsten (Halbach & Horsten, 2005) for a proof.
[5]Horsten, Leigh, Leitgeb and Welch (Horsten, Leigh, Leitgeb, & Welch, 2012) defend a similar position. According to them, the fact that FS is not naturally extensible by reflection principles as a result of ω-inconsistency means that the theory is '[...] ultimately not very attractive'.

As a consequence, supporters of FS must regard his own theory as unsound, for they fall into the following dilemma: If they commit themselves to the falsity of RFL_{FS}, which seems reasonable for $\neg RFL_{FS}$ is entailed by the system, they must embrace the unsoundness of his own theory, FS. On the other hand, if they commit to RFL_{FS}, also attractive since RFL_{FS} states the soundness of FS, supporters are forced to admit that his theory entails falsities and, hence, is unsound, for it entails the negation of RFL_{FS}.

Naturally, whatever is provable in FS it must be provable by a finite number of applications of NEC and CONEC. As a result, every theorem of FS is a theorem of an ω-consistent fragment of FS. ω-inconsistency is not the reason why FS proves the negation of its own reflection principle, but why this negation is false according to the theory itself. For while a finiteness argument runs and, thus, $\neg RFL_{FS}$ is provable in some ω-consistent subtheory of FS, RFL_{FS} only becomes true—and its negation false—when applications of NEC are unrestrained.

Nonetheless, many have argued that ω-inconsistency is harmless as long as it doesn't affect what can be proved in the language of the base theory. Horsten (Horsten, 2011, p. 158), for instance, claims that:

> Even though FS is not outright inconsistent, its ω-inconsistency seems clear evidence of its unsoundness. But things are not as bad as they seem at first sight. The *arithmetical* soundness of FS can be used to try to remove the appearance of unsoundness of FS that is created by McGee's theorem. (His italics)

In Halbach and Horsten (Halbach & Horsten, 2005, p. 216), we can find a similar point of view:

> The effects of the ω-inconsistency are limited to the sphere of the diagonal sentences involving T, where our intuitions about the notion of truth are pretty much of no use anyway.

However, ω-inconsistency *is* the cause of the fact that FS entails falsities such as $\neg RFL_{FS}$, against which their supporters are supposed to have many intuitions. In other words, even though $\neg RFL_{FS}$ is provable in some ω-consistent subtheory of FS, $\neg RFL_{FS}$ only becomes false when applications of NEC are unrestrained (because in this case RFL_{FS} is true). In fact, the theory entails $T\ulcorner A \urcorner$ if it also entails $Bew_{FS}\ulcorner A \urcorner$ for each $\mathcal{L}_T$-sentence A and, at the same time, that $\neg\forall x(Bew_{FS}(x) \to Tx)$. This is no diagonal sentence (though it is arithmetically equivalent to one, as every other $\mathcal{L}_T$-formula).

In the case of PAŁT we exhibit a different problem. As we mentioned, in this theory we have all the instances of T-schema. This is a strongly desirable feature. However, unlike FS, in this theory all the compositionality axioms fail. In fact, if we add $T\neg$ or $T \to$ as axioms, we obtain a trivial system[1]. Thus, in any non-standard model, these sentences are false because there are non-standard numbers interfering in the interpretation of the predicates. For we can prove that all instances of these principles that result from instantiating the quantifiers with (codes of) 'real' sentences are true. Yatabe (Yatabe, 2011b, p. 101) underestimates this unpleasant result:

> [...] the failure of the formalized commutation scheme [$(T\neg)$] is not enough to reject [PAŁT] on the ground that T does not satisfy the formal commutativity, therefore the failure of the formalized commutativity is not a serious problem on truth conceptions.

We do not believe that only the failure of compositional principles is enough to reject a theory of truth. Nevertheless, this is another odd consequence of ω-inconsistency. Yatabe claims that the 'semantic problem' is secondary and focuses on the high computational content of the theory. However, as we mentioned, even though this relation between

[1]See Hàjek et al. (Hàjek et al., 2000) for a proof.

complexity and ω-inconsistency in this theory can be interesting from a technical point of view, we have claimed that, from a philosophical point of view, ω-inconsistency entails too many unpleasant consequences. In what sense we can embrace a theory of truth, if the truth predicate cannot be interpreted as a genuine truth predicate? What is the meaning of predicating truth of something that is not a truth bearer? How could it be accepted that the predicate that represents the sentences of the theory has codes that do not codify sentences in its extension? Anyone who holds an ω-inconsistent theory of truth should answer these conceptual questions.

4. Conclusions

In sum, we have shown that ω-consistency seems to be a highly desirable feature for a theory of truth. Theories as FS and PAŁT do not comply with this requirement and, hence, are unable to correctly express semantic properties of $\mathcal{L}_T$. If a theory of truth is not ω-consistent, (i) it cannot be interpreted as talking about the intended ontology of arithmetic and its truth predicate; hence, (ii) it is not able to provide a better understanding of this model; (iii) the extension of the truth predicate will always contain objects that are not (codes of) sentences; (iv) also, we can't interpret the syntactic predicates as genuinely representing the syntax of the theory; and (v) it will have significant counter-intuitive consequences in the language with the truth predicate.

Regarding the point (v), FS and PAŁT lead to undesirable consequences by different reasons. On the one hand FS proves the negation of its own reflection principle, the principle that states that FS is sound. From the point of view of FS's theorists this principle must be true. However, as its negation is provable, ω-inconsistency puts the FS's supporters in the position of accepting a theory whose unsoundness cannot be avoided. On the other hand, in the case of PAŁT the commutative principles that show how the truth predicate interact with the connectives are false in every model. Non-standard numbers cause an interference in the interpretation of the truth predicate and, even though we can prove all the instances of these principles, adding the generalized versions of the principles to the theory leads to inconsistency.

Finally, we have shown that the concept of truth (and its expressive power) doesn't necessarily involve non-standardess (as Yatabe (Yatabe, 2011b) claims in his analysis of co-induction). All of this impels us to draw the following moral: a theory of truth should not only be satisfiable, but also ω-consistent.

References

anonymized. (2010). Details omitted for double-blind reviewing. -.

anonymized. (2013). Details omitted for double-blind reviewing. -.

Bacon, A. (2013). Curry's Paradox and ω-inconsistency. *Notre Dame Journal of Formal Logic*, *54*, 87–104.

Beall, J. (2009). *Spandrels of truth.* Oxford University Press.

Cobreros, P., Égré, P., Ripley, D., & Van Rooij, R. (2014). Reaching transparent truth. *Mind*, fzt110.

Field, H. (1999). Deflating the conservativeness argument. *The Journal of Philosophy*, *96*(10), 533–540.

Field, H. (2008). *Saving truth from paradox.* New York: Oxford University Press.

Field, H. (2015). Indicative conditionals, restricted quantification, and naive truth. *The Review of Symbolic Logic*, 1–28.

Fjellstad, A. (2016). Omega-inconsistency without cuts and nonstandard models. *The Australasian Journal of Logic*, *13*(5).

Friedman, H., & Sheard, M. (1987). An axiomatic approach to self–referential truth. *Annals of Pure and Applied Logic*, *33*, 1–21.

Gupta, A., & Belnap, N. D. (1993). *The revision theory of truth.* Cambridge: MIT Press.

Hàjek, P. (1998). *Metamathematics of Fuzzy Logic.* Springer.

Hàjek, P., Paris, J., & Shepherdson, J. (2000). The Liar Paradox and Fuzzy Logic. *The Journal of Symbolic Logic*, *65*, 339–346.

Halbach, V. (1994). A system of complete and consistent truth. *Notre Dame Journal of Formal Logic*, *35*, 311–327.

Halbach, V. (2011). *Axiomatic theories of truth.* Cambridge: Cambridge University Press.

Halbach, V., & Horsten, L. (2005). The deflationist's axioms for truth. In B. Armour-Garb & J. C. Beall (Eds.), *Deflationism and paradox.* Oxford University Press.

Horsten, L. (2011). *The tarskian turn: Deflationism and axiomatic truth.* Cambridge: MIT Press.

Horsten, L., Leigh, G. E., Leitgeb, H., & Welch, P. (2012). Revision revisited. *The Review of Symbolic Logic*, *5*(04), 642–664.

Kaye, R. (1991). *Models of peano arithmetic.* Oxford University Press, USA.

Ketland, J. (1999). Deflationism and tarski's paradise. *Mind*, *108*(429), 69–94.

Leitgeb, H. (2001). Theories of truth which have no standard models. *Studia Logica*, *68*(1), 69–87.

Leitgeb, H. (2007). What theories of truth should be like (but cannot be). *Philosophy Compass*, *2*(2), 276–290.

McGee, V. (1985). How truth-like can a predicate be? a negative result. *Journal of Philosphical Logic*, *14*, 399–410.

Picollo, L. M. (2013). Yablo's paradox in second-order languages: Consistency and unsatisfiability. *Studia Logica*, *101*(3), 601–617.

Restall, G. (1992). Arithmetic and Truth in Łukasiewicz's infinitely valued logic. *Logique et Analyse*, *12*, 140–303.

Robinson, A. (1963). On languages which are based on non-standard arithmetic. *Nagoya Mathematical Journal*, *22*, 83–117.

Shapiro, S. (1998). Proof and truth: Through thick and thin. *The Journal of Philosophy*, *95*(10), 493–521.

Sheard, M. (2001). Weak and strong theories of truth. *Studia Logica*, *68*, 89–101.

Tarski, A. (1935). The concept of truth in formalized languages. In *Logic, semantics, metamathematics* (pp. 152—278). Oxford: Clarendon Press.

Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and phenomenological research*, *4*(3), 341–376.

Yatabe, S. (2011a). A co-inductive language and truth degrees. In *The mathematics of language* (pp. 209–226). Springer.

Yatabe, S. (2011b). Yablo-like paradoxes and co-induction. In *New frontiers in artificial intelligence* (pp. 90–103). Springer.

Yatabe, S. (2015). Truth as a logical connective. In *Jsai international symposium on artificial intelligence* (pp. 166–182).

Appendix A.

In this Appendix, we will prove that FS and PAŁT are ω-inconsistent.

Let's start with the case of FS. Its ω-inconsistency is a consequence of McGee's (McGee, 1985) ω-inconsistency theorem[1]

Let f be a primitive recursive function defined in $\mathcal{L}_T$ and represented in PA by $\dot{f}$, such that

- $\dot{f}(0, \ulcorner A \urcorner) = \ulcorner A \urcorner$
- $\dot{f}(Sx, \ulcorner A \urcorner) = \ulcorner T\dot{f}(x, \ulcorner A \urcorner)\urcorner$

Consider the formula $\neg\forall x T\dot{f}(x, y)$ with exactly one free variable y. Then, by diagonalization, there exists a sentence C such that

$$PA \vdash C \leftrightarrow \neg\forall x T\dot{f}(x, \ulcorner C \urcorner)$$

From FS' axioms of compositionality and an application of NEC, $FS \vdash C$ and, therefore, $FS \vdash \neg\forall x T\dot{f}(x, \ulcorner C \urcorner)$. But also, by iterated applications of NEC one obtains that $FS \vdash T\ulcorner C \urcorner$, $FS \vdash T\ulcorner T\ulcorner C \urcorner\urcorner$, $FS \vdash T\ulcorner T\ulcorner T\ulcorner C \urcorner\urcorner\urcorner$, ..., which, by compositionality, entails that $FS \vdash T\dot{f}(\bar{n}, \ulcorner C \urcorner)$ for all $n \in \omega$. Consequently, FS is an ω-inconsistent theory.

This result has direct consequences for FS' possible interpretations. Since every natural number must satisfy $T\dot{f}(x, \ulcorner C \urcorner)$ but there also must be something in the domain that satisfies $\neg T\dot{f}(x, \ulcorner C \urcorner)$, only models containing more than just natural numbers in their domain are allowed. Hence, FS has only non-standard models. *A fortiori*, the FS's models restricted to the purely arithmetic fragment cannot be isomorphics with N.

As a corollary, the revision theory of nearly stable truth $T^{\#}$ developed in (Gupta & Belnap, 1993) will turn out to be ω-inconsistent too, but in a semantical sense. Usually, we call a semantical system $\mathcal{S}$ ω-inconsistent if both $\mathcal{S} \vDash \exists x \neg A$ and also $\mathcal{S} \vDash A(\bar{n})$ for each $n \in \omega$.

Let's considere now, the case of PAŁT[1]. In this case, ω-inconsistency is not a consequence of McGee's theorem[2]. Let g be a primitive recursive function defined in $\mathcal{L}_T$ and represented in PAŁT by $\dot{g}$, such that

- $\dot{g}(0, y) = y \dot{\rightarrow} \ulcorner \bot \urcorner$
- $\dot{g}(Sx, y) = y \dot{\rightarrow} \dot{g}(x, y)$

Diagonalizing the predicate $\exists x T\dot{g}(x, y)$, we get a sentence D, such that $D \leftrightarrow \exists x T\dot{g}(x, \ulcorner D \urcorner)$ is a theorem of the theory. It is easy to check that for any model M, $0 < \|\exists n T(\dot{g}(n, \ulcorner D \urcorner))\|_M \leqslant 1$. This is because if $\|D\|_M = 0$ then $\|T(\dot{g}(0, \ulcorner D \urcorner))\|_M = 1$. But then $\|\exists n T(\dot{g}(n, \ulcorner D \urcorner))\|_M = 1 \neq 0 = \|D\|_M$, a contradiction. But also, for each $n \in \omega$, $\|T(\dot{g}(\bar{n}, \ulcorner D \urcorner))\|_M \neq 1$. This can be checked looking at the clause of the conditional. But, analyzing g, $\|D\|_M = 1 = \|\exists n T(\dot{g}(n, \ulcorner D \urcorner))\|_M$. So, as in the case of FS there must be a non-standard element in the domain of every PAŁT model satisfying $T(\dot{g}(x, \ulcorner D \urcorner))$[3]. Therefore, no extension of N is a model of PAŁT.

[1] McGee (McGee, 1985) proved a general result of ω-inconsistency for a whole family of truth theories with specific features. It's application to FS is due to Friedman and Sheard (Friedman & Sheard, 1987) and Halbach (Halbach, 1994).

[1] We will not consider the differences between strong ω-inconsistency and weak ω-inconsistency, defined by Bacon (Bacon, 2013). For our purposes it's enough to show that, there is a formula $\exists x A(x)$, such that $\vdash \exists x A(x)$, even though $A(\bar{n}) \vdash$, for each, $n \in \omega$.

[2] The most common proof of the ω-inconsistency of this theory involves the so-called "modest liar paradox" (cf. Hàjek (Hàjek, 1998), Hàjek et al. (Hàjek et al., 2000), Yatabe (Yatabe, 2011b)). Here, we present another proof, based on Bacon (Bacon, 2013). However, given the features of the connectives presented for Łukasiewicz logic, both proofs are equivalent.

[3] For simplicity, here we gave a semantic proof. For a syntactic proof of this fact, see Bacon (Bacon, 2013).