Read on the liar

Read y el mentiroso

Gabriel Sandu
Institut D'histoire Et De Philosophie Des Sciences Et Des Techniques: IHPST, Francia.
University of Helsinki, Finlandia. discufilo@ucaldas.edu.co

Recibido el 30 de julio de 2011 y aprobado el 3 de octubre de 2011

Resumen

En este artículo se analiza la solución de Read al mentiroso y se señalan algunos de sus problemas con relación a otras soluciones en la literatura.

Palabras clave

Mentiroso, Parsons, Read, esquema T, teorías de la verdad.

Abstract

In this paper I analyze Read's solution to the Liar and point out some of its problems relating it to other solutions in the literature.

Key words

Liar, Parsons, Read, T-schema, theory of truth.

Tarski and the Liar

It is well known that Tarski's requirement on a materially adequate theory of truth, namely that it should entail all instances of the schema

(T) x is true if and only if p (1.)

where 'x' stands for the name of a sentence 'p', leads to paradoxes. Let us recall how the paradox is derived in Tarski's seminal paper.

We stipulate that c denotes the sentence written on the 8th line of this section:

c is false.

Applied to c the relevant instance of schema (T) is

'c is false' is true if and only if c is false.(2.)

(2) together with the previous stipulation

c = 'c is false.

entails

c is true if and only if c is false. (3.)

Finally, (3) and the principle of bivalence lead to the contradiction

c is true and c is false.

The general mechanism involved should to be clear. On one side, the language contains expressions α ('c' in our example) which denote the sentence '$\neg Tr(\alpha)$'. In addition, the language contains standard names a (' ' in our example). The relevant instance of Tarski's T-schema is

$$Tr(a) \leftrightarrow \neg \mathrm{Tr}(\alpha)$$

which together with $a = \alpha$ and the laws of identity implies a contradiction.

Truth-value gaps

Kripke's fixed point construction

One solution is to restrict the *T*-schema as Tarski did. This solution has been criticized for various reasons that will not be repeated here. One of the first serious attempts to break with the Tarskian approach was that of Kripke (1975) and Martin-Woodruff (1975).

In more details, we can take Kripke's starting point to a first-order language $\mathcal{L}$ of arithmetic which contains names for its sentences. We add a truth predicate *Tr* and form the extended language $\mathcal{L}^+ = \mathcal{L} \cup \{Tr\}$. On the interpretational level, we start with an interpretation $\mathfrak{M} = (\mathcal{U}; I; I^+; I^-)$ where $\mathcal{U}$ is the universe and *I* assign to the nonlogical vocabulary of the language appropriate elements from $\mathcal{U}$ in a standard way. The new element is the pair of functions I^+; I^- who interpret the truth predicate in a partial way: $I^+(Tr)$ is the extension of *Tr*; $I^-(Tr)$ is the counter-extension of *Tr*, disjoint from I^+ (Tr). Thus the universe $\mathcal{U}$ may be seen as divided into (a) sentences which belong to the extension $I^+(Tr)$ of the truth-predicate; (b) sentences which belong to its counter-extension; and (c) nonsentences.

The kernel of Kripke's proposal is a fixed point construction resulting in a partial model $\mathfrak{M} = (\mathcal{U}; I; E^+; E^-)$ where E^+ contains exactly the sentences true in $\mathfrak{M}$, and E^-contains exactly the sentences false in $\mathfrak{M}$. The Liar sentence '$\neg Tr(\alpha)$' is neither true nor false.

The problem with this solution is well known. If '$\neg Tr(\alpha)$' is neither true nor false, then it is not true. But then the sentence $Tr(a)$ which asserts that '$\neg Tr(\alpha)$' is true, is false. But we cannot say that consistently. For if '$Tr(a)$' is false, then by logic, '$\neg Tr(\alpha)$' is true, and thus, by one of the laws of identity, '$\neg Tr(\alpha)$' is true. But if '$\neg Tr(\alpha)$' is true, then '$Tr(a)$' which says of '$\neg Tr(\alpha)$' that it is true, should be true. Applying again the law of identity, we infer that '$Tr(a)$' is true. Thus from the premise that '$Tr(a)$' is false we ended up with the conclusion that '$Tr(a)$' is true.

Thus the Strong-Kleene proposal of Kripke which says of the Liar that it is neither truth nor false is inadequate for it does not allow one to classify the *Liar* sentence in one's own object language. That can be done only in the metalanguage where one has availbale the notion of contradictory negation.

IF-logic

Another attempt to overcome Tarski's second impossibility result is given within the so-called *IF*-languages introduced by Hintikka and Sandu (1989) (see Hintikka 1996 Hodges 1997). These languages express more quantifier dependencies and independencies than ordinary first-order languages whose extensions they are. More concretely, the object language contains sentences of the form

$$\forall x_0(x_1/\{x_0\})\ (\exists x_2/\{\ x_1\})\ (\exists x_3/\{\ x_0,\ x_2\})\ R(x_0;\ x_1;\ x_2;\ x_3) \qquad (4.)$$

which are meant to express the fact that:

$\forall x_1$ is not in the scope of $\forall x_0$;
$\exists x_2$ is not in the scope of $\forall x_1$;
$\exists x_3$ is not in the scope of $\forall x_0$ nor in that of $\exists x_2$.

The slash is thus an outscoping device. The sentence (4) expresses in a linear notation the so-called Henkin or branching quantifier introduced by Henkin (1961):

$$\begin{pmatrix} \forall x_0\ \exists x_2 \\ \forall x_1\ \exists x_2 \end{pmatrix} R(x_0;\ x_1;\ x_2;\ x_3) \quad (5.)$$

The truth-conditions of (5) (and alternatively (4)) are given by a translation in a second-order metalanguage. The basic ideas for this translation are those of game theoretical semantics (GTS). With every φ the object-language $\mathfrak{L}$, model $\mathfrak{M}$ for $\mathfrak{L}$ and assignment s (which is empty in case $\square$ is a sentence) a semantical game is associated, $G\ (\varphi;\ \mathfrak{M};\ s)$, opposing *Eloise* (the initial verifier) to *Abelard* (the initial falsifier). In the relevant game associated with (4), the players choose alternatively the elements a; b; c and d from the universe of $\mathfrak{M}$ to be the interpretations of the four quantifiers (Both players have two choices corresponding to the universal and respectively the existential quantifiers). The play (a; b; c; d) is a win for *Eloise* if it belongs to the interpretation $R^{\mathfrak{M}}$ of R in $\mathfrak{M}$. Otherwise it is a win for *Abelard*. The slash codes the information sets of the players in the semantical games. Thus for her first move *Eloise* knows only *Abelard's* first choice, and for the second move she knows only *Abelard's* second choice. The sentence (4) is true in the model $\mathfrak{M}$ (relatively to the assignmend s) if and only if there is a winning strategy for *Eloise* in the game $G\ (\varphi;\ \mathfrak{M};\ g)$, that is, there are two functions f; g

defined only on the possible known moves so that <a; b; $f(a)$; $g(b)$> is a win for *Eloise* for any a and b chosen by *Abelard*. And similarly (4) is false in $\mathfrak{M}$ if and only if there is a winning strategy for *Abelard*, that is, there are elements x and y such that <x; y; c; d; g> is a win for *Abelard* for any choices c and d of *Eloise*. The truth ($\mathfrak{M} \models^+$) and falsity ($\mathfrak{M} \models^-$) of (4) is given by the following translations (we abbreviate $\forall x_0(x_1/\{x_0\})(\exists x_2/\{x_1\})(\exists x_3/\{x_0, x_2\})$ by $Hx_0x_1x_2x_3$):

$$\mathfrak{M} =^+ Hx_0x_1x_2x_3R(x_0, x_1, x_2, x_3) \Longleftrightarrow \exists f \exists g \forall x_0 \forall x_1\ R(x_0, x_1, f(x_0)\ g(x_1))$$
$$\mathfrak{M} \models^+ Hx_0x_1x_2x_3R(x_0, x_1, x_2, x_3) \Longleftrightarrow \exists x_0 \exists x_1 \forall x_2 x_3 \neg R(x_0, x_1, x_2, x_3)$$

There are *IF*-sentences in the pure language of identity (e.g. $\forall x_0(\exists x_1/\{x_0\}))$ which are neither true nor false in any model which contains at least two elements. It was shown in Sandu (1996 1998), Hyttinen and Sandu (2000) that there is an *IF*-formula $\Psi(x)$ in the vocabulary of *PA* which defines "true-in-$\mathfrak{M}$" for every model $\mathfrak{M}$ of *PA*; that is

$$\mathfrak{M} \models^+ \varphi \Longleftrightarrow \mathfrak{M} \models^+ \Psi(\ulcorner \varphi \urcorner)$$

for every *IF*-sentence φ in the vocabulary of *PA*.

As in the previous case, the Liar which now takes the form of a sentence β denoting '$\neg\Psi(\beta)$' is neither true nor false (the negation $\neg$ is interpreted as role switching). One is not able to express this fact in the object language but only in an extension containing classical negation. That negation, however, cannot be any longer interpreted game-theoretically, an interesting fact by itself which cannot be discussed here.

Common to both Kripke's partially interpreted languages and *IF*-languages is the fact that $\varphi \rightarrow \varphi$ is defined as $\neg\varphi \vee \varphi$. Accordingly, for sentences φ like the Liar which are neither true nor false, the implication $\varphi \rightarrow \varphi$ is not valid in Kripke's semantics, neither in *IF*-logic.

In the paper "The truth schema and the Liar" Stephen Read offers a thought provoking solution to the *Liar via* a detour through the notion of what a sentence says. According to his view, the *Liar* turns out to be false, and this can be consistently asserted in the object language. In what follows I am going to question this solution. Before doing it, I will give a short presentation of Parsons' solution to the *Liar* which, I think, offer an interesting point of comparison with that of Read.

Parsons: quantifier shift

One way to get around the problem discussed in connection with truth-valued gaps has been suggested by Charles Parsons (1983) and Tyler Burge (1979). Their idea is that the truth-predicate as it appears in the Liar applies to different entities than the truth-predicate which is used to classify the Liar sentences. There is one essential modification though with respect to Tarski's theory: it is propositions and not sentences which are the truth-bearers. Consider now the reformulation of the Liar sentence in terms of propositions

(c) c expresses a false proposition.

In order to allow for the possibility of a sentence not expressing a proposition at all, Parsons replaces Tarski's T-schema by the weaker:

$\forall x$(x is a proposition and 'p' expresses x, then x is true if and only if p). (6.)

Together with the assumption that propositions are bivalent, (6) entails

$\forall x$(x is a proposition and 'p' expresses x, then x is false if and only if $\neg p$). (7.)

Applied to the two Liar sentence, (6) and (7) lead to the conclusion that they do not express a proposition at all. Here is the argument:

Suppose x is a proposition and c expresses x. Then by (6) we get

x is true if and only if c expresses a false proposition. (8)

Suppose x is not true. By existential generalization we infer

$\exists x$(x is a a proposition $\wedge \neg$(x is true) $\wedge$ c expresses x) (9.)

that is, c expresses a false proposition. But now, from (8) we get that x is true. Thus starting from an arbitrary proposition x that c expresses, we landed in the conclusion that x is true.

$\forall x$((x is a proposition $\wedge$ c expresses x) $\rightarrow$ x is true) (10.)

(10) is equivalent with

$$\neg\exists x(x \text{ is a a proposition} \wedge \neg(x \text{ is true}) \wedge c \text{ expresses } x). \quad (11.)$$

But then there is no proposition that *c* expresses, for if *c* expressed one, say *y*, then by (10) *y* would have to be a true proposition, and that together with (8) implies that *c* expresses a false proposition. But this is in contradiction with (11). A similar argument shows that the Strenghthened Liar (*d* is the sentence: *d* does not express a proposition) does not express a proposition either.

As Parson notices, there is a difficulty with his proposal: (c) says of a certain sentence which is (c) itself, that it expresses a false proposition. The argument above has shown that (c) does not express any proposition. But then (c) seems to say something false. Aren't we compelled to say that (c) expresses a false proposition after all? If the answer is yes, it can be shown that a contradiction will arise, and we end up in the same predicament in which we found ourselves with the truth-value gaps solution: the Liar sentence cannot be classified within the object language.

Parsons avoids the contradiction by his shifting quantifier domain assumption. According to it, both '*c* expresses a false proposition' and '*c* does not express any proposition' are true, but the quantifiers range over a domain of propositions which is different from the propositional domain relevant for assessing (c) in the first place. The former is larger than the latter.

The quantifier-shift proposal has been found attractive for the possibility it opens up for narrowing narrow down the gap between set-theoretic and semantic paradoxes introduced by Ramsey. Here is Parsons' argument. Given a predicate '*Fx*', the fact that a is its extension is expressed by the condition

$$\forall x(x \in a \leftrightarrow Fx). \quad (12.)$$

By analogy with (6) above we have

$$\forall y(y \text{ is the extension of } 'Fx' \rightarrow \forall x(x \in y \leftrightarrow Fx)). \quad (13.)$$

If we now take '*Fx*' to be '$x \notin$ 'x, we obtain

$\neg\exists y\forall x(x\in y \leftrightarrow x \notin x)$. (14.)

But (13) and (14) entail

$\neg\exists y$(y is the extension of ′$x \notin$′x). (15.)

Parsons adopts here the same solution he proposed for the Liar, that is, he takes the two quantifiers in (13) and (15) to range over distinct domains (Parsons 231-232).

Read: the Liar is false

Read takes seriously what Tarski thought at a certain moment to be the philosophical motivation for his theory of truth, namely a correspondence theory encoded in the principle

(CP) A sentence is true if and only if things are as the sentence says they are.

Read abbreviates ′x says that p′ by ′$x : p$′ where ′x′ designates a sentence. The notion of ‘saying that’ is a technical notion, a close relative to Frege’s notion of content in the *Begriffsschrift* according to which the content of a sentence comprises all its logical consequences. Read does not explicitly draw the analogy with Frege, but he nevertheless wants his notion of *saying that* to be closed under the principle

(***K***) $\forall p,q(p \Rightarrow q) \rightarrow (x : p \Rightarrow x : q)$

Where ′$\Rightarrow$′ is strict implication and ′$\rightarrow$′ is material implication. It is clear that a sentence says, in this technical sense, more than what it says in the intuitive sense. The crux of Read’s proposal is to replace Tarski’s *T*-schema by

$T(x) \Leftrightarrow \forall p(x : p \rightarrow p)$. (A)

In other words:

(S) x is true if and only if things are wholly as x says they are.

Read is aware that (A) makes true all sentences which says nothing and is thus in need of qualification. The way he qualifies it is to conjoin $\exists p(x : p)$

to the right-hand side of (A). For reasons of simplicity, this proposal is not followed but Read assumes, instead, that each sentence to which (A) is applied says something. I shall return to this point later on.

According to Read, the point of replacing the (T)-schema with the (A)-schema is that, unlike the former, all the instances of the latter are true. In this new setting, the liar sentence c turns out to be false without contradiction, and, amazingly, the laws of classical logic still hold. The argument is resumed below.

The liar sentence c says that $\neg Tr(c)$. It may say more, say, $\neg Tr(c) \wedge q$. This together with (A) entails

$$Tr(c) \Leftrightarrow \forall p(c : p \rightarrow p). \qquad (16.)$$

But given that c says that $\neg Tr(d) \wedge q$ and that is all that c says, we get from (16)

$$Tr(c) \Leftrightarrow \neg Tr(c) \wedge q. \qquad (17.)$$

Thus

$$\neg Tr(c) \Rightarrow \neg\ (\neg Tr(c) \wedge q). \qquad (18.)$$

which is equivalent with

$$\neg Tr(c) \Rightarrow Tr(c) \vee \neg q. \qquad (19.)$$

(19) and (K) entail

$$c : Tr(c) \vee \neg q \qquad (20.)$$

which in conjunction with $c : q$ yields

$$c : (Tr(c) \vee \neg q) \wedge q \qquad (21.)$$

whence

$$c : (Tr(c). \qquad (22.)$$

The argument has showed that if c says that $\neg Tr(c)$ it also says that $Tr(c)$ as well, i.e. $c : \neg Tr(c) \wedge Tr(c)$. Thus by (A)

$$Tr(c) \Leftrightarrow (\neg Tr(c) \wedge Tr(c)\ldots) \tag{23.}$$

whence

$$\neg Tr(c). \tag{24.}$$

The liar sentence is thus not true. Read's conclusion is the following:

> [c] cannot be true, for to be true, it would have to be both true and not true. Nothing can be both true and not true. So c cannot be true… The solution is ready to hand. Abandon (T) and realize that the correct theory of truth is given by (A) …governing all well-formed sentences in a semantically closed language. As applied to c, we obtain the correct truth-condition:
>
> $Tr(c) \Leftrightarrow (\neg Tr(c) \wedge Tr(c))$. (Read 10)

Read's view has some close analogy with Parson's when the later is reformulated to apply to sentences. Parsons himself de.nes such an explicit truth-predicate by

$$Tr(y) \leftrightarrow \exists x(x \text{ is a a proposition} \wedge (x \text{ is true}) \wedge y \text{ expresses } x) \tag{(*)}$$

and then points out that together with (6) it implies

$$\exists x(x \text{ is a a proposition} \wedge {'p'} \text{ expresses } x) \rightarrow (Tr('p') \leftrightarrow p). \tag{(T*)}$$

Obviously, for sentences $'p'$ which do not express a proposition, we will not be able to assert the consequent of (T*).

Analogously, for falsity, he ends up with

$$\exists x\ (x \text{ is a a proposition} \wedge {'p'} \text{ expresses } x) \rightarrow (F('p') \leftrightarrow \neg p). \tag{(F*)}$$

Let us represent propositions by second-order propositional variables. Then (*) becomes
$Tr(x) \Longleftrightarrow \exists p(p$ is a proposition $\wedge\, x$ expresses $p \wedge p)$. (25.)

Analogously, we may define falsity by:

$F(x) \Longleftrightarrow \exists p(p$ is a proposition $\wedge\, x$ expresses $p \wedge \neg p)$. (26.)

Making explicit Read's abbreviation '$x : p$', his schema (A) becomes:

$Tr(x) \Leftrightarrow \forall p(p$ is a proposition $\wedge\, x$ says that $p \rightarrow p)$. (27.)

The analogy between (25) and (27) is now straightforward:

> For Parsons, a sentence is true exactly when it expresses a proposition which is the case.
> For Read, a sentence is true exactly when everything it says is the case.

Unlike Parsons, Reads claims that the *Liar* sentences express propositions (i.e., say something) and have truth-values. The key ingredient in his system which allows him to avoid a contradiction is his technical notion of "saying that".

Recall Parsons' analysis of the *Liar*: an argument has shown that the Liar sentence c does not express a proposition at all. Hence the Liar sentence asserts something false. But this is what the *Liar* sentence c says; hence it is true after all. In the end, the *Liar* sentence which is neither true nor false, receives a determinate truth-value when the domain of the quantifiers shift.

Read's solution is different. He has produced an argument showing that the *Liar* sentence c is not true (or false). One may now be tempted to adopt the same line of reasoning as above and continue

> "But this is what the *Liar* says, hence the *Liar* is true after all".

If this could be done, a contradiction would be derived. The point of the present solution is that one cannot continue in the way just described. The particularity of Read's approach is that the *Liar* does not only say that it is not true, it also says that it is true. Therefore, in order for the

Liar to be true, *everything* the *Liar* says must be the case. But as Read puts it, “nothing can be both true and false”.

I find two problems with this solution (apart from quantifications over propositions, etc).

The minor problem is the way it deals with falsity. Recalling Read’s provision devised to block his schema (A) to apply to sentences which say nothing, (27) should be rephrased as:

$$Tr(x) \Leftrightarrow \exists p(p \text{ is a proposition} \wedge x \text{ says that } p) \wedge \\ \forall p(p \text{ is a proposition} \wedge x \text{ says that } p \rightarrow p). \quad (28.)$$

Reads does not deal with falsity explicitly, but given that he accepts the principle of bivalence, (28) entails:

$$F(x) \Leftrightarrow \forall p(p \text{ is a proposition} \rightarrow \neg(x \text{ says that } p)) \vee \\ \exists p(p \text{ is a proposition} \wedge x \text{ says that } p \wedge \neg p) \quad (28)$$

That is, a sentence x is false if it either does not say anything or it says something that is not the case. In other words, all sentences which say nothing are false. This conclusion, although philosophically defensible, is absurd in my opinion, but I am not going to dwell upon it.

The feature in Read’s treatment that concerns me here is the way his notion of “saying that” is applied in the proof. He starts from the instance

$$Tr(c) \Leftrightarrow \forall p\ (c : p \rightarrow p)$$

and then from the assumption that $\neg Tr(c) \wedge q$ is all that c says, he derives

$$Tr(c) \Leftrightarrow \neg Tr(c) \wedge q.$$

But why should we assume that everything the *Liar* says is expressible by one single proposition? In fact, we should not, given the fact that the *Liar* says of it both that it is true and that it is false. Hence by the principle (K), for every p, the *Liar* says that p. In other words, there are infinitely many propositions expressed by the *Liar*. Accordingly, (23) cannot be a formula of the relevant object language (unless this language is infinitary): when we explicitate the dots on the right side of the equivalence, we obtain an infinite conjunction. For this reason, Read's argument to the effect that the *Liar* is false but not true, can be properly carried out only in a (n infinitary) metalanguage in which propositions like the ones expressed by the *Liar* can be represented. This is the price he has to pay for the acceptance of the (*A*)-scheme and of the principle (*K*). Parsons can attribute a a determinate truth-value to the *Liar* only after shifting the domain of propositions expressible by it. Similarly, Kripke can classify the *Liar* only in the metalanguage in which one has available contradictory negation, and the same goes for *IF*-logic. In Read's case, the proposition expressed by the *Liar* can be shown to receive a determinate truth-value by appeal to an argument which, when properly expressed, requires an infinitary language. The possibility of assigning a determinate truth-value to the Liar while sticking to the rules of classical logic has, in the end, turned out to be illusory.

BIBLIOGRAPHICAL REFERENCES

Hintikka, J. *The principles of mathematics revisited*. Cambridge: Cambridge University Press, 1996. Print.

Hintikka, J. & Sandu, G. "Informational independence as a semantical phenomenon". J. Fenstad et al. (Eds.). *Logic, methodology and philosophy of science VIII*. Amsterdam: Elsevier Science Publishers B.V., 1989. Print.

---. "Game-theoretical semantics". J. van Benthem & A. ter Meulen. (Eds.). *Handbook of logic and language*. Amsterdam: Elsevier, 1997. Print.

Hodges, W. "Compositional semantics for a language with imperfect information". *Journal of the IGPL 5*. 1997: 539-563. Oxford. Print.

Hyttinen, T. & Sandu, G. "Henkin quantifiers and the definability of Truth". *Journal of Philosophical Logic 29*, 2000: 507-527. Springer. Print.

Kripke, S. "Outline of a theory of truth". *Journal of Philosophy 72*. 1975: 690-716. Columbia University. Print.

Martin, R. L. (Ed.). *Recent essays on truth and the Liar paradox*. Oxford: Oxford University Press, 1984. Print.

Martin, R. L. & Woodruff, P. W. "On representing 'true-in-L' in L". *Philosophia 5*. 1975: 213-217. Springer. Print.

Parsons, Ch. *Mathematics in philosophy*. New York: Cornell University Press, 1983. Print.

Read, S. L. “The Truth Schema and the Liar”. S. Rahman, T. Tulenheimo, & E. Genot. (Eds.). *Unity truth and the Liar.* Amsterdam: Springer, 2008a. Print.

Sandu, G. “IF first-order Logic, Kripke, and 3-Valued logic”. Appendix to Hintikka. J. Hintikka. *The principles of mathematics revisited*. Cambridge: Cambridge University Press, 1996. Print.

---. “IF-logic and truth-definition”. *Journal of Philosophical Logic 27*. 1998: 143-164. Columbia University. Print.

Tarski, A. “The concept of truth in formalized languages”. A. Tarski and J. Corcoran. (Eds). *Logic, semantics, metamathematics*. Indianapolis: Hackett, 1983. Print.