SORITES (ΣΩΡΙΤΗΣ), ISSN 1135-1349

Issue #11. December 1999. Pp. 15-23.

Was Frege Wrong when Identifying Reference with Truth-Value?

WAS FREGE WRONG WHEN IDENTIFYING REFERENCE WITH TRUTH-VALUE?

by Jean-Yves Béziau

0. Introduction

Frege’s thesis according to which the reference (*Bedeutung)* of a sentence (*Satz*) is a truth-value (*Wahrheitswert*)[1] is one of the most controversial aspect of his philosophy. Many people think that it is against common sense, according to which the reference of a sentence is a state of affair, a fact.

Some people have even tried to show that Frege was absolutely wrong. This is the case of G.Sengupta in a two pages paper (Sengupta 1983) which seems to have been approved by Dummet (cf. the footnote attached to the last sentence of the paper). Here are the first two sentences of the paper:[2]

> A fundamental assumption in Frege’s semantics (henceforth A1) is that the customary reference of a declarative sentence is its truth-value. The purpose of this paper is to prove that A1 is false.

One of the main difficulties in discussing this kind of things is that although Frege’s work is the origin of many basic concepts of modern logic, they have been seriously transformed. A typical example is the Fregean stroke (). It is difficult to know exactly what was its exact meaning for Frege (in fact Frege changed several times of idea) but one thing is sure when we interpret «I — *P*» as meaning «*P* is logically true», we are using a conceptual framework which is quite different from Frege’s one although Frege’s work can be considered as its source.

However self-incoherent interpretations cannot be used against Frege, and it seems that Sengupta’s argumentation is based on such an interpretation.

---

[1] We will not discuss here terminological problems and we will stick to these conventional translations. For a discussion about Frege’s notion of *Bedeutung*, see e.g. (Angelelli 1982).

[2] Our analysis will lead us to quote most of Sengupta’s short paper, so that it is not necessary to read it in order to understand Sengupta’s argumentation and our refutation of it.

After stating various possible interpretations of Frege's principle of substitution (section 1), we show that there is no coherent interpretation under which Sengupta's argumentation is valid (section 2). Finally we try to see how Frege's distinction can work in the context of modern mathematics and how modern logic grasps it (section 3).

1. Substitution

Let us first quote a fundamental extract of Frege's *Sinn und Bedeutung* where he justifies in a sense his option of identifying reference with truth-value:

> Wenn unsere Vermutung richtig ist, dass die Bedeutung eines Satzes sein Wahrheitswert ist, so muss dieser unverändert bleiben, wenn ein Satzteil durch einen Ausdruck von derselben Bedeutung, aber anderm Sinne ersetz wird. Und das ist in der Tat der Fall. Leibniz erklärt gradezu: *Eadem sunt, quae sibi mutuo substitui possunt, salva veritate.* Was sonst als der Wahrheitswert könnte auch gefunden werden, das ganz allgemein zu jedem Satze gehört, bei dem überhaupt die Bedeutung der Bestandteile in Betracht kommt, was bei einer Ersetzung der angegebener Art unverändert bliebe ? (Frege 1892, p.35)[3]

According to Frege, we can therefore state the following substitution principle, which he sees himself as an interpreation of Leibniz's principle:

Frege's substitution principle

It two sentences *Q* and *Q'* have the same truth-value, thus the sentence *P* containing *Q* as a subsentence has the same truth-value as the sentence *P'* that we get from *P* substituting *Q'* for *Q*.[4]

Within the framework of present mathematical logic, this principle can be interpreted in two differents ways: on the one hand taking Fregean truth to be simple truth (truth in a model), on the other hand to be logical truth. Accordingly there are two definitions of substitutions which are not equivalent.

---

3 If our supposition that the reference of a sentence is its truth value is correct, the latter must remain unchanged when a part of the sentence is replaced by an expression having the same reference. And this is in fact the case. Leibniz gives the definition: «*Eadem sunt, quae sibi mutuo substitui, salva veritate*'. What else but the truth value could be found, that belongs quite generally to every sentence if the reference of its components is relevant, and remains unchanged by substitutions of the kind in question ? (Max Black's translation).

4 This is already an interpretation of Frege, in fact an adaptation to the case where the Satzteil is itself a *Satz*.

Substitution's principle 1 (S1)

It two sentences Q and *Q'* have the same truth-value in a given world (or model, or valuation),[5] thus the sentence *P* containing *Q* as a subsentence has the same truth-value (in this world) as the sentence *P'* that we get from *P* substituting *Q'* for *Q*.

Substitution's principle 2 (S2)

It two sentences *Q* and *Q'* are logically equivalent, i.e. are true in exactly the same worlds, thus the sentence *P* containing *Q* as a subsentence is logically equivalent, i.e. is true in exactly the same worlds, to the sentence *P'* that we get from *P* substituting *Q* ' for *Q*.

As we will see in the next section, the mistake of Sengupta is due to the

fact that he interprets Frege's principle as an incoherent mixture of (S1) and (S2).

We will now make a few remarks about these principles to clarify their meanings and in order to give a basis for the analysis presented in our third section.

First let us note that in modern logic the term *substitution* is used in several different ways. In general by the law (rule or theorem) of substitution it is meant something which neither corresponds to (S1) nor (S2), but the fact that if, in a tautology, we substitute a given sentence for all the occurrences of an atomic sentence, it is still a tautology.

(S2) is generally called the *replacement* theorem (e.g. Kleene's terminology) although it is also sometimes presented under the name *substitution* theorem (e.g. Church's terminology).

(S1) itself rarely appears under such a name. This principle is most of the time not stated explictly. It is obviously true in any matrix's semantics. In particular if we say that a logic is *truth-functional* iff it can be characterized by a finite matrix,[6] (S1) holds in every truth-functional semantics. The validity of (S1) in matrix's semantics is due to the fact that in this case the truth-value (under a given valuation) of a compound sentence is a function of the truth-values of its components.

It is possible to prove that (S2) holds in every truth-functional bivalent logic (i.e. logic which can be charcaterized by a two-valued matrix and therefore for which (S1) holds); see e.g. (Béziau 95). It is a consequence of the fact that from the viewpoint of a two-valued matrix, we can replace in (S2) «is true» by «have the same truth-value», i.e. in this case S2 is equivalent to the following principle:

---

5 We are not precise in order to include the widest range of semantics (sentential, first-order, Kripke, etc.)

6 We will stick to this definition, which seems to be the implicit one when someone says that modal logics or intuitionistic logic are not truth-functional. For a discussion about this question, see (Béziau 1997).

Substitution's principle 3 (S3)

If two sentences *Q* and *Q'* have the same truth-values in exactly the same worlds, then the sentence *P* containing *Q* has the same truth-values in exactly the same worlds as the sentence *P'* that we get from *P* substituting *Q'* for *Q*.

It is clear that there are some logics in which (S2) holds but not (S1). For example if we consider the current modal logics, from the point of view of Kripke's semantics, (S2) holds but not (S1).[7]

2. Refutation of Sengupta's proof

These definitions being made, let us turn to Sengupta's interpretation and argumentation:

> We shall take for granted the verity of the assumption that the truth-value of a declarative sentence is a function of the references of its parts (henceforth A2). A2 is not only in conformity with Frege's view, but also entailed by Leibniz's principle. A consequence of A1 and A2 is that the truth-value of a declarative sentence containing another as part remains unchanged when the part is replaced by another sentence having the same truth-value, provided that the part as part has only customary reference and expresses a complete thought. Since we have taken the verity of A2 for granted, if the consequence is proved to be false so is A1.

Let us call A3 what Sengupta calls «a consequence» of A1 and A2, deleting the final part which is in fact independent of Sengupta's mistake. Thus we have the following assertions:

> A1. The customary reference of a declarative sentence is its truth-value.
>
> A2. The truth-value of a declarative sentence is a function of the references of its parts.
>
> A3. The truth-value of a declarative sentence containing another as part remains unchanged when the part is replaced by another sentence having the same truth-value.

A3 looks very much like Frege's substitution principle. In particular the question if we must interpret it as (S1) or as (S2) is left open. However we can remark that Sengupta articulates A1, A2 and A3 in a particular way. To take A3 as a consequence of A1 and A2 seems to choose to interpret A3 as (S1). It is not obvious that this articulation corresponds to Frege's one.

[7] Łukasiewicz's three-valued logic, in which both (S1) and (S2) hold, was supposed to formalized the notion of possibility, but nowadays nobody considers this logic as a modal logic.

Sengupta gives the following description of the example, according to which he will (allegedly) prove that A3 is false and that therefore, A2 being assumed, Frege cannot claim A1:

> Let us consider the following sentences assuming that Srimati (…) detests long hair:
>
> 1. Two plus two is equal to four
>
> 2. Srimati detests long hair.
>
> 3. It is unfortunate for Ranjan that Srimati detests long hair.
>
> (…)

Sentence 1 is necessarily true and under the assumed circumstances 2 is also true.

We can thus say that Sengupta chooses a world *w* («the assumed cicrumstances») in which 2 is true and 1 also, because according to him the sentence 1 is true in all the worlds («is necessarily true»). Let us note that the sentence 1 is not very well chosen in the sense that the fact that it is a necessary truth is controversial. It would be better to take a tautology like:

1’. If Srimati detests long hair then Srimati detests long hair.

Then Sengupta goes on as follows:

> Now, if Frege were right in assuming that the customary reference of a declarative sentence is its truth-value, then 1 and 2 would be coreferential, and substituting the one for the other in sentence 3 would have no consequence for its truth-value, provided that the embedded sentence in sentence 3 had only customary reference and expressed a complete thought.

After showing that 3 had only customary reference and expresses a complete thought (parts of the argumentation which is of no interest for us here), Sengupta concludes his paper as follows:

> The consequence of substituting 1 for 2 in sentence 3 remains to be seen.The substitution does not necessarily preserve the truth-value of sentence 3. We can easily conceive of possible worlds in which the fact that Srimati detests long hair is unfortunate for Ranjan, but not the fact that two plus two is equal to four. A1 is thus proved to be false.

What Sengupta is saying is that there are some worlds in which 3 is true and the following sentence 3’ is false:

3’. It is unfortunate for Ranjan that two plus two is equal to four.

But what can we conclude from that ? All we can say is that 3 and 3’ are not logically equivalent. But 1 and 2 are not logically equivalent. Thus this does not contradict (S2).

Imagine now that Sengupta consider that the consequence A3 of A1 and A2 is (S1) and not (S2). How can he say that (S1) is false, and that assuming A2, therefore he has proved that A1 is false ?

In the given world *w*, taking A2 for granted, 3 and 3' should have the same truth-value,[8] since in *w* 1 and 2 have the same truth-value. Thus (S1) is not contradicted.

Therefore the consequence A3 of A1 and A2, should it be (S1) or (S2), is not proved to be false.

In fact it seems that Sengupta in order to refutate Frege is using the following principle of substitution, which is an absurd mixture of (S1) and (S2) that no one would defend:

Sengupta's substitution principle

If two sentences *Q* and *Q'* have the same truth-value in a given world, thus the sentence *P* containing *Q* as a subsentence is true in exactly the same worlds as the sentence *P'* that we get from *P* substituting *Q'* for *Q*.

3. Reference as class of models

We will now try to show how Frege's distinction can be articulated within the framework of mathematics and how modern logic captures it. This account will shed a new light on the relations between truth-functionality, extensionality and intensionality.

Most people identify truth-functionality with extensionality, and therefore, taking intensionality as the opposite of extensionality, they identitfy non-truth-functionality with intensionality. According to these views, current modal logics are intensional because they are not truth-functional.

Our proposal leads us to think that extensionality is expressed by (S2) and that it differs from truth-functionality (only bivalent truth-functionality entails extensionality in the sense of (S2), as remarked in the first section). In particular current modal logics are extensional (because (S2) holds) even if they are not truth-functional.

The solution of the identity paradox within present mathematical logic and the construction of a real intensional logic seems therefore open problems.[9]

Let us consider the axioms for complemented distibutive lattices, in short CDL. We can say that the reference (*Bedeutung*) of these axioms is the class of their models. That is to say, following Tarski's idea, the class of structures in which they are true. This same class can be given in many other ways, that is to say, with different sets of axioms. For example the axioms IR for idempotent rings.

---

[8] We will not discuss here the question if it is appropriate to think that A2 applied to the sentence 3. This is what Sengupta assumes and assumes that Frege assumed.

[9] For more details on this question see (Béziau 1994).

The fact that CDL and IR refer to the same thing, the class of boolean algebras, is not necessarily evident. This was proved by Marshall Stone after a tedious conceptual work and was a fundamental step for the proof of his famous representation theorem, cf. (MacLane 1981). This result was an important discovery of the same kind as the discovery that Hesperus and Phosphorus refer to the same object.

A boolean algebra can be seen as a complemented distributive lattice or as an idempotent ring, these are two different ways of looking at the same object. CDL and IR are two different manners of having access to one and the same thing. They are two different *meanings* for the same reference, according to Frege saying that the meaning (*Sinn*) is *the way of giving* (*die Art des Gegebenseins*) the reference (*Bedeutung*), cf. (Frege 1892, p.26).

Using the extension/intension terminology, we can say that CDL and IR are two different intensions for the same extension.[10]

Because the replacement theorem is valid in classical first-order logic, formulas (or set of formulas) having the same extension, can be identified (the relation of logical equivalence is a congruence). This is what happens with CDL and IR formalized in the context of classical first-order logic.

In fact classical first-order logic minimizes the rôle of meaning, interpreted along the above lines, and is not able to give an account to it.

From the viewpoint of the mathematician, the difference of meanings between CDL and IR appears relatively clearly: CDL is formulated in the language of order and IR in the language of function, these two languages corresponding to two different basic intuitions. Of course it is a rough distinction and no mathematician has given yet a precise definition which supports such kind of theory of meaning.[11] But it seems that it fits Frege's view according to which the meaning (*Sinn*) of a sentence (*Satz*) is a thought (*Gedanke*). We can say that what the mathematician feels and tries to explain is that CDL and IR are two different ways of thinking (at the same thing).

In first-order logic the difference of these two languages is very tiny. In fact within first-order logic what is emphasized is the possibility of reduction (modulo the replacement theorem): for example, functions can be defined as predicates.[12]

Within the framework of a (classical first-order) modal logic in which the replacement theorem holds (which is the case of current modal logics), the two following sentences are equivalent:

Stone proved that a complemented distributive lattice is a boolean ring.

---

[10] Here we have the following equation: intension = comprehension = axiomatization. The axioms are the comprehensive way of giving the extension, i.e. the class of models.

[11] However Bourbaki's description of mathematics (as it appears in Bourbaki 1950) gives a key to such a theory.

[12] But model theory reaches great achievements through this line, showing that if we succeed to express a class of structures in a particular way, it reveals important properties of it.

Stone proved that a boolean ring is a boolean ring.

Therefore modal logics do not solve the identity paradox. According to them Stone, like George IV, is the son of La Palice. And the reason why is that they are purely extensional and are not able to express the distinction between reference (*Bedeutung*) and meaning (*Sinn*).[13]

Bibliography

Angelelli, Ignacio. 1982. «Frege's notion of 'Bedeutung'». *Logic, Methodology and Philosophy of Science VI*, L.J.Cohen *et al.* (eds), North Holland, Amsterdam, pp.735-753.

Béziau, Jean-Yves. 1994. «Du Pont's paradox and intensional logic». *Logica'93 (Proceedings of the 7th International Symposium)*, P.Kolár and V.Svoboda (eds.), Filosofia, Prague, pp.62-65.

Béziau, Jean-Yves. 1995. *Recherches sur la logique universelle*. PhD. Denis Diderot University (Paris 7).

Béziau, Jean-Yves. 1997. «What is many-valued logic ?» *Proceedings of the 27th Symposium on Multiple-Valued Logics*, IEEE Computer Society, Los Alamitos, pp.117-121.

Bourbaki, Nicolas. 1950. «The architecture of mathematics». *American Mathematical Monthly*, **57**, pp.221-232.

Frege, Gottlob. 1892. «Über Sinn und Bedeutung», *Zeitschrift für Philosophie und philosophische Kritik*, **100**, pp.25-50.

Gabbay, Dov M. and F.Günthner (eds.). 1986. *Handbook of philosophical logic, vol.II*, Reidel, Dordrecht.

MacLane, Saunders. 1981. «History of abstract algebra: origin, rise and decline of a movement». *American mathematical heritage: algebra and applied mathematics*, D.Tarwater *et al.* eds, Texas University, pp.l1-35.

Sengupta, Gautam. 1983. «On identifying reference with truth-value». *Analysis*, **43**, pp.72-74.

Jean-Yves Béziau

Universidade Federal Fluminense. Instituto de Matemática

Niterio, Rio de Janeiro, Brazil

<**jybeziau@hotmail.com**>

---

[13] In the above example, we can put instead of «Stone proved»: «Stone believed», «Stone thought», «It is necessary», «It is possible», etc, all these expressions being considered as interpretations of operators of current modal logics or similar logics (epistemic logics, provability logic, etc.). The fact that in the literature these logics are quite commonly presented as intensional logics is typically illustrated by the cover of the second volume of the *Handbook of Philosophical Logic*: it is said that it «surveys the most significant «intensional» extensions (sic) of predicate logic» (Gabbay 1986).