Self-referencing languages revisited

Gábor Rédey[1] and Attila Neumann[2]

[1] Hungarian Atomic Energy Authority
H-1036 Budapest, Fényes A. u. 4., Hungary
[2] Technische Universität Wien
Institut für Computergraphik und Algorithmen
Favoritenstraße 9-11 / E186, A-1040 Wien, Austria

Abstract. Paradoxes, particularly Tarski's liar paradox, represent an ongoing challenge have long attracted special interest. There have been numerous attempts to give either a formal or a more realistic resolution to this area based on natural logical intuition or common sense.
The present semantic analysis of the problem components concludes that the traditional language of logic fails to detect Tarski's paradox, since the formalised version of a liar sentence does not represent a correct definition. Neither the formal language, nor the logical system is deficient in this respect. Only natural language statements cannot be interpreted adequately by traditional language of logic.

1 Introduction

Paradoxes play a remarkable role in philosophy and logic. Several have been resolved, eliminated or brought to rest by appropriate theories (e.g.: Cretan paradox: Epimenides the Cretan says "All Cretans are liars", or the key statement of Nihilism: "there is no truth"), while others attract perpetual and changeless interest. Time and again new researches they challenge, and consistently elude resolution.

One of the most impressive paradoxes is analysed in this paper, specifically Tarski's liar paradox associated with the definition of logical truth. Much effort has been made since Tarski's theory of meta languages to obtain a semantically more acceptable explanation. For a thorough, complete and even historical analysis of this matter the reader can be referred to Feferman [2], although, there are many other papers continuously re-examining this point [1, 4, 5, 3].

Self-referencing languages with surely imply a fundamental dilemma within philosophical logic. However, this property also demands particular interest from the Artificial Intelligence community. Clearly, a potential attribute of representation languages is beneficial in reasoning systems of any kind, no matter whether it is for a human or robot. Natural languages are essentially rich in introspective statements concerning feelings, remarks, opinions, knowledge or other content-related features. Assertions of this kind apply a large variety of linguistic devices such as indirect quotations or modalities which are also occasionally loaded by self-reference. Therefore, an adequate representation of self-reference is crucial not only theoretically or philosophically, but also from a practical, functional point of view.

The problem occurs not at the level of natural languages, but that of the formalisation process. Formalisation of some natural language sentences can be inadequate even if it is possible. The latter alternative can obviously be ignored, since it would impugn the possibility of natural human understanding, without which there is no meaningful communication, formal or otherwise. The justification and resolution of the former option is the objective of this paper.

The foremost exact phrasing and modelling of the informal liar sentences was originated by Tarski and subsequently by numerous alternatives by Kripke, Gilmore, Feferman, Perlis, Kerber and others. It will be shown that the initial formal rephrasing of the liar sentence does not satisfy the natural requirement of being a valid definition.

The paper presents an attempt to answer this problem, which is based on a different logical language, determined by one of the authors, which has been proven to have a better expressive power.

2 Traditional formalising self-reference

The problem itself cannot even be identified without any formal representation. The first realization relating a Cretan type of sentence is accomplished by Tarski.

2.1 Tarski's paradox

Roots of Tarski's paradox are summarised here from [2]. Let $x, \ldots$ range over the statements of the language of a logical system, which are assumed to be closed under the usual propositional operators denoted by $\sim$, $\&$, $\vee$, $\supset$, $\equiv$. Each statement x of the language has a name, i.e. there is an associated closed term $\ulcorner x \urcorner$ of the language. Then the following axiom is accepted for a predicate $t(\ulcorner x \urcorner)$, which is interpreted as expressing that x is true:

$$t(\ulcorner x \urcorner) \equiv x \tag{1}$$

for each statement x of the language.

For the derivation of a contradiction in this system the liar sentence is taken:

$$\text{(2) } \textit{is not true.} \tag{2}$$

formalised mostly as

$$l :\equiv \sim t(\ulcorner l \urcorner) \tag{3}$$

By definition (1) and transitivity of biconditional we obtain $t(\ulcorner l \urcorner) \equiv \sim t(\ulcorner l \urcorner)$, immediately implying inconsistency of the system.

2.2 The solution routes

There has been considerable work on theories eliminating these paradoxes. Numerous, seemingly different escape routes have been determined the common idea of which is to discard problematic sentences:

1. either by altering the syntax so that undesirable statements could be excluded from the language,
2. or by revising critical axioms semantically so that antinomic formulae could be evaluated exceptionally (e.g. as meaningless ones).

Because of limited space, individual instances of the above categories are not discussed here. The former is followed by Tarski and Kerber [3], while the latter is preferred by Feferman [2] and Perlis [4, 5].

3 An examination of formal liar sentences

As was shown previously, naive truth theory is considered to be destroyed by paradoxes of self-reference, inspiring to create alternative theories avoiding these antinomies. The following approach tries to review the representation technique of the liar sentence challenging naive truth theory.

3.1 Liar sentence translated by biconditional

Many authors discussing the liar paradox represent sentence (2) as was shown in section 2.1, e.g. [2, 4]. Obviously,

$$l \equiv \sim t(\ulcorner l \urcorner) \tag{4}$$

is not a literal translation of (2), thus many authors disagree with this formula.

Another reason against the use of the biconditional here is that according to first-order logic the next logical consequence holds:

$$(p \equiv q) \vdash\dashv \sim(p \equiv \sim q) \tag{5}$$

Interpreting this for (1) yields

$$x \equiv t(\ulcorner x \urcorner) \vdash \sim (x \equiv \sim t(\ulcorner x \urcorner)) \tag{6}$$

that makes the defining scheme of the liar sentence *false* immediately. In other words this translation of the liar sentence cannot cause any paradox relative to truth definition.

3.2 Liar sentence translated by equation

The previous, (4), interpretation of the liar sentence (2) is rather controversial. Another common representation of (2) is the following:

$$\ulcorner l \urcorner = \ulcorner \sim t(\ulcorner l \urcorner) \urcorner \tag{7}$$

The only difference between (7) and (4) is that the biconditional is replaced by an equation. This change makes the translation of (2) clearer, although one may still have doubts concerning it, as the consequences remain the same.

At this point contradiction can be deduced, if $\ulcorner l \urcorner$ is substituted into (1), then Leibniz's rule and transitivity of biconditional is applied:

$$t(\ulcorner l \urcorner) \equiv l \equiv \sim t(\ulcorner l \urcorner) \tag{8}$$

According to the accepted reasoning this step concludes the commonly known paradox.

Nevertheless, the above deduction suffers from the same defects that in the previous section. It is easy to show by the truth definition (1) together with Leibniz's rule, $(a = b) \supset (F(a) \supset F(b))$, that

$$(\ulcorner x \urcorner = \ulcorner y \urcorner) \supset (x \equiv y) \tag{9}$$

from which by the substitution $\sim t(\ulcorner x \urcorner)/y$ and applying the contraposition rule

$$\sim (x \equiv \sim t(\ulcorner x \urcorner)) \supset \sim (\ulcorner x \urcorner = \ulcorner \sim t(\ulcorner x \urcorner) \urcorner) \tag{10}$$

is reached. Then from this latter and (6)

$$\sim (\ulcorner x \urcorner = \ulcorner \sim t(\ulcorner x \urcorner) \urcorner) \tag{11}$$

can be deduced, that contradicts to the assumption (7).

Now it is shown that even the modified representation of the liar sentence, (7), cannot cause paradox relative to the truth definition (1), as the scheme $\ulcorner x \urcorner = \ulcorner \sim t(\ulcorner x \urcorner) \urcorner$ is also evaluated *false*.

Essentially, this means that translations of (2), i.e. (4) and (7) hitherto discussed, are unsound as definitions, and thus fail to represent (2) adequately.

4 Finding a way out

Section 3 has shown that the root problem with liar type self-referencing is at an earlier level than was expected. The formula scheme working as the definition of a liar sentence is unravelled as a *false* scheme. This fact seems to weaken the commonly known formal proofs of liar paradox, hence an apparent question probing the grounds of this indefinite phenomenon arises. On the other hand, this antinomy is still present in natural languages [8]. Thus, it cannot be due to a defect in natural language. There is no choice but to assume that the representation itself is invalid.

The question remains open whether a logical type of representation which is able to render this content adequately can exist. The rest of this paper argues that an isomorphic representation is necessary for this purpose, as described in the next section.

5 Representing a liar by iCTRL

The previous sections have shown that traditional language of logic has difficulties in representing self-referential sentences in their natural form. The imperfect fidelity of translation may be a sufficient cause of improper interpretation of the case. Thus, changing the logical representation language may effect a more adequate model of the phenomenon. We present an alternative and novel manner of representing the matter relying on intensional conformal text representation language (iCTRL) initiated by one of the authors [6, 7]. It is a knowledge representation tool closer to natural languages, preserving not only truth as traditional logical language does, but it also models natural grammatical relations. Accordingly, it seems to be suitable for a better formalisation of self-reference, at the same time it allows the verification of soundness of the preceding issues.

There is only limited space here to give a detailed formal introduction to iCTRL, so the reader is referred to [6, 7]. Although, a rather reduced sublanguage of iCTRL is sufficient for the whole description, the reader should be made familiar with some new notation. We focus immediately on truth definition. The previously discussed definition (1) can be now written as

$$t\ x, \langle\alpha\rangle x. \equiv \alpha \tag{12}$$

where α represents any arbitrary proposition, $\langle\alpha\rangle x$ the name of this proposition, α, while $t\ x$ stands for the truth predicate, and $t\ x, \langle\alpha\rangle x.$ expresses the statement that α is a true statement.

This latter expression deserves more attention. The predicate symbol $t\ x$ itself appears to be the same as in classical logic, but its application to the corresponding name symbol differs from the way as it is commonly treated: $t(\langle\alpha\rangle)$, since name symbols are treated now as singular predicate symbols, i.e. they must have an argument. Finally "," and "." are punctuation symbols, the former links the sentence predicate and its subject together, providing they share a common variable symbol, the latter just closes the sentence. Evaluation of the actual sentence $t\ x, \langle\alpha\rangle x.$ simply answers the expectations: it is true if and only if the extension of $\langle\alpha\rangle x$ is completely included in the extension of $t\ x$.

Before trying to represent self-reference, the ordinary case should be presented briefly. Regularly, subjects, like $\langle\alpha\rangle x$, and corresponding predicates, such as $t\ x$, share the same variable symbol, because they are referring to the same group of individuals. However, an exterior subject, mentioned earlier or later in a text, cannot be referred to in this way, only by a compound reference variable term. E.g. $y : x$ can provide that compound term redirecting its left side variable parameter to its right side one that refers to that subject expression sharing the same variable parameter name. Let the pair of sentences *John walks. He whistles.* be considered. The corresponding pair of formulae in this context is $walk\ x, John\ x.\ \ whistle\ y : x$. A referred subject can naturally be eliminated by inserting it as it is referred to: $walk\ x, John\ x.\ \ whistle\ y, John\ y.$[3]

[3] As a matter of fact, subject reference needs a bit more complex notation, which was simplified here to reduce unacquainted formalism to the minimum.

After these preliminary notes the iCTRL formula which exactly formalises the liar sentence (2) is

$$\langle \sim t\ x : y \rangle y. \tag{13}$$

Now the sentence itself plays the role of the subject of the predicate $t\ x$, the variable parameter of which x is redirected to the referred subject by $t\ x : y$.

Considering (12), it does not appear to significantly differ from the earlier version of truth definition. Nevertheless, the liar sentence representative (13) is quite dissimilar to the classical formulae, (4) and (7), respectively. According to the construction, it is a literal translation of (2). It does not comprise any extraneous constituents such as biconditional or equation, which are also auxiliary tools in the corresponding classical formulae.

Contradiction results from a substitution of $\langle \sim t\ x : y \rangle y.$ into truth definition (12), that is $\sim t\ x, \langle \alpha \rangle x. \equiv \sim \alpha$ generating:

$$\sim \langle \sim t\ x : y \rangle y. \tag{14}$$

Accomplishment of this substitution appears to be strange, because the liar sentence predicate $\sim t\ x : y$ inside, wrapped into the subject part of the sentence, is to be matched by the left side of (12), $\sim t\ x, \langle \alpha \rangle x$.

The other pair of contradictory statements, similar to (14) and (13), correspondingly causes a paradox. If negation of liar $\sim \langle \sim t\ x : y \rangle y.$, is substituted into (12) that is $\sim (\sim t\ x, \langle \alpha \rangle x). \equiv \alpha$, that similarly yields $\langle \sim t\ x : y \rangle y.$, then that is a contradiction.

In conclusion, the liar sentence has been proven plainly to be antinomic showing that inconsistency based on this kind of argumentation is clearly achieved. Traditional attempts to explain this make the impression that natural language and formal language of logic have been drifted apart. However, this is not the case for iCTRL.

6 Closing Remarks

The source of Tarski's semantical paradox has been revised in this paper concluding with the recognition that a liar sentence, which is traditionally applied to generate an explicit antinomy with the classical truth definition, fails to give an effective argument against the related conventional extension of first-order logic. This conclusion can be deduced in each formalisation instance of the classical language of logic originating from the fact that the liar sentence definition fails to define the liar sentence itself. iCTRL modelling enabling a formal syntactic fidelity of translation from natural languages, can prove this paradox case exactly. The approach presented here has shown an adequate representation of self-reference that may stimulate further development with respect to representation techniques of introspection.

Acknowledgements

The core idea of this paper was emerged during a colloquium with Prof. Weissmahr. The authors would like to express their indebtedness to him for his invaluable help and encouragement without which this work could never have come to fruition.

The authors would like to thank the anonymous reviewers for their help that made this paper more concise and distinct.

References

1. Gilmore, P. C.: The consistency of partial set theory without extensionality, *Proc. of Symposia in Pure Mathematics*, **13/II** (1974) 147–153.
2. Feferman, S.: Toward useful type-free theories. I., *J of Symbolic Logic*, **49/1** (1984) 75–111.
3. Kerber, M: On Knowledge, Strings and Paradoxes, Proceedings of JELIA, Springer Verlag, 1998.
4. Perlis, D.: Languages with self-reference I: foundations, *Artificial Intelligence*, **25** (1985) 301–322.
5. Perlis, D.: Meta in Logic, in Maes, P.; Nardi, D. eds.: Meta-level architectures and reflection, Elsevier Science B.V., 1988, 37–49.
6. Rédey G.: Conformal Text Representation, *Engineering Applications of AI*, **6** (1993) 65–71.
7. Rédey G.: iCTRL: Intensional Conformal Text Representation Language, *Artificial Intelligence*, **109/1-2** (1999) 33–70.
8. Sainsbury, R.M.: Paradoxes, second edition, Cambridge University Press, 1995.