How to Represent Opaque Sentences in First Order Logic

Bijan Arbab
IBM Corporation
Los Angeles Scientific Center
2525 Colorado Ave. 3rd floor
Santa Monica, California 90406

Abstract

This paper presents a method for applying standard inferencing mechanisms to a broader class of sentences than that which was possible before. The logic of *proposition surrogates* allows representation of and reasoning with a class of sentences (the so called opaque sentences) that pose special difficulties for standard logics. Within this class are sentences with one or more occurrences of such words as *know, believe, aware, search, hunt, etc.*

It is shown that standard formal (programming) languages, e.g. first order logic, can be extended with proposition surrogates to deal with facts that have traditionally been expressed in modal or various other proposed logics. It has been argued that such facts can not adequately be expressed in standard logics; the findings and results recorded here, however, are to the contrary. Proposition surrogates can be added, in a conservative manner, to standard automatic reasoning systems.

Proposition surrogates and their historical development are presented. An inference engine based on the logic of proposition surrogates is then outlined and applied to some problems in this area.

1 Introduction

In *Be griff sschrift,* a formula language for pure thought modeled upon that of arithmetic, Gottlob Frege (1879) takes identity to be a relation between names or signs of objects:

$$\vdash (A \equiv B)$$

mean that *the sign A and the sign B have the same conceptual content, so that we can everywhere put B for A and conversely.*

This is the basic idea behind the axioms of equality which are assumed by various systems of logic. Following is an explicit list of the axioms of equality: reflexivity $x = x$, symmetry $x = y \rightarrow y = x$, transitivity $x = y \wedge y = z \rightarrow x = z$, substitution in predicates

$$x_i = x_0 \wedge p(x_1, \ldots, x_i, \ldots, x_n) \rightarrow p(x_1, \ldots, x_0, \ldots, x_n)$$

substitution in functions

$$x_i = x_0 \rightarrow f(x_1, \ldots, x_i, \ldots, x_n) = f(x_1, \ldots, x_0, \ldots, x_n)$$

A number of theorem provers[1] that are based on such logics implement the axioms of equality either directly or by various methods and rules of inference. It is because of the axioms of equality that certain conclusions can follow from premises. For example if it is true that eight is greater than five

$$8 > 5 \qquad (1)$$

and that the atomic number of oxygen is eight

$$f(o) = 8 \qquad (2)$$

it can then be concluded, on the bases of 1, 2 and the axioms of equality, that:

$$f(o) > 5 \qquad (3)$$

which truthfully expresses the proposition that the atomic number of oxygen is greater than five. In the paper titled *On Sense and Denotation,* Frege (1892) expresses dissatisfaction with his earlier choice of the identity relation. He explaines in detail why a name cannot always be replaced by another of the same truth-value or content, in the view of the invariance of the truth of the whole sentence. It is of course assumed that declarative sentences denote a truth value (either true or false) and express a proposition (the objective content which is capable of being the common property of many), just as names have a denotation (the particular object named) and a sense (the manner and context of presentation). Ajdukiewicz (1967) illustrates the same point with the following example:

If it is true that Newton knew that eight is greater than five

$$knew(Newton, 8 > 5) \qquad (4)$$

then it can be concluded, on the basis of 2, 4 and the axioms of equality that:

$$knew(Newton, f(o) > 5) \qquad (5)$$

which is certainly not true since it expresses the proposition that Newton knew that the atomic number of oxygen is greater than five (a fact which was beyond his ken).

[1]For example, see the work of Wos and Robinson (1969) and more recently Digricoli and Harrison (1986).

How is it, then, that a *sound* system of logic admits false conclusions based on true premises and standard rules of inference?

According to the terminology of Church (1983), this problem is called *the parados? of the name relation.* A number of radically different solutions have been proposed to solve the paradox of the name relation. Following are various contrasting views whence the source of the problem lies and how it should be solved.

2 Philosophical Views

2.1 Sense and Denotation

Frege's (1892) solution to this paradox revolves around the idea that names, sentences, or signs have associated with them a sense (the proposition expressed) which is no less relevant than the denotation. He also identifies three different contexts, *ordinary, direct,* and *indirect,* in which names can be used. In an ordinary context, names have their customary denotation and sense. The direct context is what is now known as the use-mention distinction: words name (denote) other words[3]. In an indirect context, names denote their customary sense, not their customary denotation, and have an indirect sense which is different than their customary sense.

The paradox is resolved since formula 4 is about the customary sense of the number eight, not its customary denotation, and formula 2 is about the customary denotation of the number eight, not its customary sense. Therefore, formula 2 does not warrant the substitution of *f{o)* for 8 in formula 4. Frege did not present a formulation, similar to that provided in *Begriffsschrift,* for the logic of sense and denotation.

Church presents three different alternatives under which a formulation of the logic of sense and denotation can be carried out. The three alternatives-Alt (2), *Alt(1), Alt(0}~* correspond to different sets of assumptions under which two sentences can be considered to have the same sense or express the same proposition. That two sentences *S* and 51 have the same sense if and only if *S ~ S1* is *logically valid* is called *Alt(S).* A stronger criterion of identity between senses, *Alt(l),* is that *S* is convertible to 51 according to the rules of lambda calculus. The strongest criterion of identity between senses, *Alt(O),* is that 5 and 51 differ at most by one or more alphabetic changes of bound variable, or one or more interchanges of synonymous notations. Two names are synonymous if they have the same denotation as well as the same sense.

A sound system of axioms characterizing two of these alternatives, *Alt(2) and Alt(l),* has been specified by Church (1973, 1987). McCarthy (1979) also presented a first order theory of individual concepts and propositions based on Frege's solution[4]. Formulation of the logic of sense and denotation remains open under *Alt(O),* the strongest alternative under which two sentences can be considered to express the same proposition. The solution presented in this paper is under *Alt(O),* however, it differs from the logic of sense and denotation.

2.2 Contextual Descriptions

Russell's (1905)[5] solution to the paradox eliminates names altogether from the language, and introduces *contextual descriptions.* The relevant distinction is that contextual descriptions have no meaning of their own; however, every sentence in which they occur has a meaning. It was commonly believed that the theory of contextual description can be used to resolve the paradox of the name relation as well as other paradoxes. By providing counter examples Church (1983) demonstrates that, if intensionality is to be avoided, then the theory of contextual descriptions cannot be adopted as a solution to the paradox of the name relation. Contextual descriptions, however, remain useful for solving a variety of other problems.

2.S Nonclassical Logics

The notion of possible worlds has recently received a lot of attention from philosophers because it can be used to provide an analysis of necessity and possibility. More recently it has also been applied to propositional attitudes such as believing and knowing. A number of different modal logics based on the possible world models have been proposed.

There are disagreements, however, among philosophers regarding the nature of these possible worlds. Some say that possible worlds combine the actual world with other worlds that contain only things similar to those in the actual world. Others say that a possible world is described by a set of propositions, such that each proposition or its negation is a member of the set. Some of the modal logics based on possible world semantics unnecessarily commit the agents to be what Hintikka (1975) called *logically omniscient*[6]. The strongest objection to nonclassical logics is the lack of efficient inferencing mechanisms. Construction of efficient inference engines for modal logics must also address the computational complexities of logics that are based on the possible world models.

2.4 Proposition Surrogates

This paper presents a modification, Arbab (1988), of the solution first proposed by Ajdukiewicz (1960) and later formalized by Church (1983). The solution follows the

[2]Carnap (1956) used the word *antinomy,* but the word *paradox* is preferable since no apparent contradiction occurs in the absence of any further assumptions.

[3]In writing, quotation marks or *italics* are used for direct contexts.

[4] It differs, however, from Frege's solution in that the latter calls for an infinite hiearchy of senses where as the former allows only a single level. An infinite array of senses is called for since various levels of indirection (Pat knows that Newton knew that ...) can easily be formed.

[5] In 1903, Russell had outlined a different solution to the paradox of the name relation. Russell (1905), however, flatly states that the Russell (1903) solution is very similar to Frege (1892), and both are shown to be unsatisfactory. The particular line of reasoning presented by Russell (1905) remains unclear to this author!

[6]An exception to this is Church's (1951) formulation of the logic of and sense and denotation under *Alt(2)* which is also based on the possible world models.

philosophy that there is nothing inherently wrong with the underlying logic (either the rules of inference or the axioms of equality); therefore, it is unnecessary to construct new logics, e.g., modal logics, or to abandon (or weaken) the axioms of equality. The source of the problem lies in how natural language sentences are to be formulated in the formal language. In short, if well-formed formulas corresponding to English sentences are written correctly, then paradoxical conclusions will not arise. For example, 4 is not the correct representation of the fact that Newton knew that eight is greater than five, since it leads to paradoxical conclusions. What are, then, the well-formed formulas corresponding to natural language sentences?

The answer is based on the idea that sentences denote truth-values and express propositions. The proposition expressed by a sentence is, of course, independent of the particular natural language in which it happens to be written and can be expressed by sentences in different languages. For example, the two sentences *eight is greater than five,* in English, and *Acht ist groBer als funf,* in German, both express the proposition that eight is greater than five. One method of encoding propositions within a formal language is presented in this paper. The encodings are called *proposition surrogates,* since in the formal language they play the role of the proposition expressed by a sentence. The algorithm for constructing proposition surrogates can then be added to any formal language. In this paper an inference engine based on the proposition surrogate solution is presented and applied to some examples.

3 Solution

Ajdukiewicz (1960, 1967) argues that if sentences are in an indirect context, then they are ambiguous. It is primarily this ambiguity that leads to paradoxical conclusions. There are at least two different meanings that can be attributed to such sentences. Let us call these Ψ and Ω. The paradox arises because we understand the sentence on the basis of Ψ, but formulate it on the basis of Ω. The solution, then, is to formulate the sentence according to the understood meaning. For example, the first meaning, Ψ, of the sentence

Newton knew that
the atomic number of oxygen is greater than five (6)

can informally be stated as follows:

Newton knew
about the atomic number of elements,
about oxygen,
about the relation greater than,
about the number five,
that the atomic number of oxygen is
greater than five.

Sentence 6 can be formulated on the basis of $ as follows:

$$knew(Newton, < \lambda F\lambda A\lambda G\lambda B < G, < F, A >, B >, f, o, gt, 5 >) \quad (7)$$

The existence of propositions are not effected by the fact that they can or can not be expressed in a particular natural or formal language.

where /, o, *gt,* and 5 correspond, respectively, to the atomic number of elements, the constant *oxygen,* the relation *greater than,* and the constant *five.* The ordered n-tuple occurring in the second position of the *knew* predicate is a proposition surrogate. Its first member corresponds to the form of the formula, and the rest contain the primitive constants, function, and predicate symbols that occur in the formula. For every *about* clause in Ψ, there is an appropriate constant, function, or predicate name in the proposition surrogate.

The formula which corresponds to a proposition surrogate can be obtained by applying the first member of the proposition surrogate which is always a lambda expression containing the particular *form* of the formula to the rest of the members which are always the particular *constants* occurring in the formula. The formula which corresponds to the above proposition surrogate is *gt(f(o),* 5), which expresses the proposition that the atomic number of oxygen is greater than five.

The second meaning, Ω, of 6 can informally be stated as follows:

Newton knew
about the atomic number of oxygen,
about the relation greater than,
about the number five,
that the atomic number of oxygen is
greater than five.

Sentence 6 can be formulated on the basis of Ω as follows:

$$knew(Newton, < \lambda A\lambda G\lambda B < G, A, B >, f(o), gt, 5 >) \quad (8)$$

The distinction between the two meanings can now be made clear by examining the different number of *about* clauses. According to Ψ, Newton knew about the atomic number of elements [/], oxygen [o], and that the value of the atomic number of oxygen [/(o)] is greater than five. If sentence 6 is understood according to Ψ, then its truth value is falsehood, since knowledge of the atomic number of elements [/] can not be attributed to Newton. Sentence 6 understood according to *Ω,* however, is true. The second meaning, *Ω,* of 6 does not attribute explicit knowledge of the atomic number of elements [/] or oxygen *[o]* to Newton: they do not individually appear among the primitive constants of the proposition surrogate.

The difficulty with this solution to the paradox of the name relation lies in the way primitive constants of the formalized language are handled: an equality relation between primitive constants of the formalized language can be used to reintroduce the paradox. Consider, for example, the two sentences:

John believed that Dr Jekyll was a gentleman (9)

Dr Jekyll is Mr Hyde (10)

Sentence 9 is formalized under Alt(O) as:

$$believed(john, < \lambda F\lambda A < F, A >, gent, dj >) \quad (11)$$

where *John, gent,* and *dj* are primitive constants of the language corresponding to John, gentleman, and the Dr

Jckyll respectively. Sentence 10, is formalized in the usual way as:

$$dj = mh \quad (12)$$

In formula 11, it is possible to substitute *rah* for *dj* on the basis of the axiom of equality and formula 12, thus arriving at the conclusion:

$$believed(john, < \lambda F \lambda A < F, A >, gent, mh >) \quad (13)$$

which corresponds to the paradoxical conclusion:

John believed that Mr Hyde was a gentleman (14)

An extensional solution to the problem of primitive constants is outlined below. The idea of pointers to constants (address of a particular cell within the memory of a computer) and the associated operators (obtaining the address and de-referencing) is well-known in the field of computer science. The analogy[8] between pointers and what Frege (1892) called the sense of a name can be used to construct a solution to the problem of primitive constants of proposition surrogates.

Church's 1983 algorithm for obtaining the proposition surrogate under Alt(0) is modified so that every occurrence of a primitive constant, say c, in the proposition surrogate is replaced by @c (the particular pointer to c); and every occurrence of a bound variable in the body of the lambda term corresponding to the primitive constant c, is replaced by application of the + (de-referencing) operator to that variable. For example, 9 will now be formalized under Alt(0) as:

$$believed(john, < \lambda F \lambda A < *F, *A >, @gent, @dj >) \quad (15)$$

and although *dj* — *rnh,* it does not follow that $@dj$ — $@mh$. Thus, the paradoxical conclusion:

$$believed(john, < \lambda F \lambda A < *F, *A >, @gent, @mh >) \quad (16)$$

is avoided. Additionally, the proposition surrogate in 15 may be reduced to $< *(@gent), *(@dj) >$ which in turn can be reduced to $gent(*(@dj))$ and finally to $gent(dj)$.

4 Example

Let us assume, then, the availability of a two place predicate *ps(SyP)* such that *P* is the proposition surrogate corresponding to *S* under Alt(0), for the exact details and the Prolog code see Arbab (1988). This example is called the *The Wise Man Puzzle* and has been used to test the representational ability of formalisms for knowledge representation.

> *A certain King wishes to determine which of his three wise men is the wisest. He arranges them in a circle so that they can see and hear each other and tells them that he will put a white or black spot on each of their foreheads, hut that at least one spot will be white. In fact, all three spots are white. He then offers his favor to the one who will first tell him the color of his spot. After a while, the wisest announces that his spot is white. How does he know?*[9]

The solution to this puzzle requires a wise man to reason about what other wise men know and do not know, from observations and the king's announcements. The puzzle solved here is actually a simplified version of the original puzzle. The simplifying assumptions arc that there are only two wise men, and that after some time the first wise man announces that he cannot tell the color of his spot, whereupon the second wise man says his own spot is white. The following is a partial list of a formulation of the puzzle in first-order logic, and Prolog modified with proposition surrogates (the complete list can be found in Arbab (1988)).

- That each wise man knows that there is at least one white spot is expressed by

  $$(x)(y)x \neq y \supset know(x, \lfloor \neg W(x) \supset W(y) \rfloor) \quad (17)$$

  Note that | P J is short hand for the proposition surrogate of P under Alt(0) obtained according to the algorithm presented in the last section. The corresponding Prolog clause is:

```
know(X, P) :-
      ps(w(Y) :- non(w(X)), P),
      diff(X, Y).
```

  Note that *ps(S,* P) is true iff *P* is the proposition surrogate of *S.* Also, if *P* is unbound then the Prolog interpreter will compute the proposition surrogate of *S* and bind it to *P.*

  That each wise man knows that the other wise man knows that there is at least one white spot, is expressed by

  $$(x)(y)x \neq y \supset know(x, \lfloor know(y, \lfloor \neg W(x) \supset W(y) \rfloor) \rfloor) \quad (18)$$

  The corresponding Prolog clause is:

```
know(X, P1) :-
      ps(w(Y) :- non(w(X)), P),
      ps(know(Y, P), P1),
      diff(X, Y).
```

  Notice the doubly-indirect context- thus the two calls to the *ps* predicate.

- That *A* does not know that he has a white spot, since he made an announcement to this effect, is expressed by

  $$\neg know(A, \lfloor W(A) \rfloor) \quad (19)$$

  The corresponding Prolog clause is:

  non(know(a, w(a))).

- That *D* knows that A does not know that he has a white spot, since *B* heard A's announcement, is expressed by

  $$know(B, \lfloor \neg know(A, \lfloor W(A) \rfloor) \rfloor) \quad (20)$$

  The corresponding Prolog clause is:

[8]This is not to suggest that Frege's (1892) sense of a name is simply a pointer to that name, only that pointers can play the *logical role* of the sense of a name.

[9]This puzzle has been attributed, by Konolige (1986), to an unpublished note by McCarthy.

```
know(b, P) :-
      ps(w(a), P1),
      ps(non(know(a, P1)), P).
```

Once these are consulted by a Prolog interpreter, the following questions can be answered according to the standard rules of inference:

- Does B know that he has a white spot? The Prolog formulation is:

  `?- ps(w(b), P),know(b, P).`

  Note that, in order to answer the above question, the system must reason not only about what B knows, but also about what B knows that A knows, and so on. The reported answer is: *yes.*

- Does B know that A has a white spot? The Prolog formulation is:

  `?- ps(w(a), P),know(b, P).`

  The reported answer is: *yes.*

- Who knows that he has a white spot? The Prolog formulation is:

  `?- ps(w(X), P),know(X, P),write(X).`

  The distinction between this question and the first one is in the use of a Prolog variable. The reported answer is: *b yes.*

- List everyone who knows that he has a white spot. The Prolog formulation is:

  `?- ps(w(X), P),know(X, P),write(X),fail.`

  The distinction between this question and the former is that the final *fail* clause will force the interpreter to look for all possible answers, to show that no incorrect solution is found by the interpreter. The reported answer is: *b no.*

5 Conclusion

It was shown that the common interpretation of the identity relation, i.e., a relation involving only denotation of names, leads to the paradox of the name relation. This is the source of inconsistencies when a classical set of axioms and rules of inference are assumed by the formal language. The philosophical point of view defended here is that the elimination of the paradox of the name relation requires neither a modification of the classical set of axioms, nor of the rules of inference.

An inference engine based on the logic of proposition surrogates can play an important role in the field of machine intelligence, for if a machine is ever to interact intelligently with other agents, machines or humans, then it must be able to represent and reason with facts about the agents' state of mind. An agent's state of mind, of course, includes, but is not limited to, facts about its knowledge, beliefs, awareness, and expectations.

The logic of proposition surrogates can be used to represent and reason with such facts. An inference engine based on the logic of proposition surrogates enables a machine not only to represent, but also to discover, logical consequences of such facts.

6 Acknowledgement

This work was carried out at UCLA under the supervision of Professor Stott Parker of the Computer Science Department, and Professor Alonzo Church of the Philosophy and Mathematics Departments. The author is grateful for their contributions of ideas and corrections made over the years. Needless to say, all remaining errors are attributable solely to the author.

References

[1] Kazimierz Ajdukicwicz. Intensional expressions. In Jerzy Giedymin, editor, *Kaztmierz Ajdukicwicz The scientific world-perspective and other essays 1931-1963,* pages 320-347, D. Reidel Publishing Company, 1967.

[2] Kazimierz Ajdukicwicz. A method of eliminating intensional sentences and sentential formulae. In *Atti del XII Congresso Internazional di Filosofia,* pages 17-24, 1960.

[3] Bijan Arbab. *A Formal Language for Representation of and Reasoning about Indirect Context.* PhD thesis, University of California at Los Angeles, 1988.

[4] Rudolf Carnap. *Meaning and Necessity.* The University of Chicago Press, second edition, 1956.

[5] Alonzo Church. Intensionality and the paradox of the name relation. In *Themes from Kaplan,* Stanford Press, 1988. The content of this paper was presented as an invited lecture at a joint symposium of the A.P.S. and the Association for Symbolic Logic in Berkeley, California, March 1983.

[6] V. J. Digricoli and M. C. Harrison. Equality-based binary resolution. *Journal of the ACM,* 33(2):253-289, 1986.

[7] Gottlob Prcge. Begriffsschrift, a formal language, modeled upon that of arithmatic, for pure thought. In Jean van Heijenoort, editor, *From Frege to Godel,* pages 1-82, Harvard University Press, 1977. originally published in 1879.

[8] Gottlob Frege. On sense and meaning. In A.P.Martinich, editor, *The Philosophy of Language,* pages 212-220, Oxford University Press, 1985. originally published in 1892.

[9] J. Hintikka. Impossible possible worlds vindicated. *Journal of Philosophical Logic,* 4:475-484, 1975.

[10] Kurt Konolige. *A Deduction Model of Belief.* Morgan Kaufmann Publishers, 1986.

[11] John McCarthy. First order theories of individual concepts and propositions. In B. Meltzer and D. Michie, editors, *Machine Intelligence 9,* pages 120-147, Ellis Horwood, 1979.

[12] Bertrand Russell. On denoting. *Mind,* 14:479-493, 1905.

[13] Larry A. Wos and G. A. Robinson. Paramodulation and theorem proving in 1st order theories with equality. In B. Meltzer and D. Michie, editors, *Machine Intelligence 4*, pages 135-150, Edinburgh University Press, Edinburgh, Scotland, 1969.