THE REPRESENTATION OF INCONSISTENT INFORMATION IN A DYNAMIC MODEL-THEORETIC SEMANTICS

Douglas B. Moran
Department of Computer Science
Oregon State University
Corvallis, Oregon 97331

ABSTRACT

Model-theoretic semantics provides a computationally attractive means of representing the semantics of natural language. However, the models used in this formalism are static and are usually infinite. Dynamic models are incomplete models that include only the information needed for an application and to which information can be added. Dynamic models are basically approximations of larger conventional models, but differ is several interesting ways.

The difference discussed here is the possibility of inconsistent information being included in the model. If a computation causes the model to expand, the result of that computation may be different than the result of performing that same computation with respect to the newly expanded model (i.e. the result is inconsistent with the information currently in the dynamic model). Mechanisms are introduced to eliminate these local (temporary) inconsistencies, but the most natural mechanism can introduce permanent inconsistencies in the information contained in the dynamic model. These inconsistencies are similar to those that people have in their knowledge and beliefs. The mechanism presented is shown to be related to both the intensional isomorphism and impossible worlds approaches to this problem.

I. INTRODUCTION

In model-theoretic semantics, the semantics of a sentence is represented with a logical formula, and its meaning is the result of evaluating that formula with respect to a logical model. The model-theoretic semantics used here is that given in The proper treatment of quantification in ordinary English (PTQ) [Montague 1973], but the problems and results discussed here apply to similar systems and theories.

From the viewpoint of natural language understanding, the conventional model-theoretic semantics used in descriptive theories has two basic problems: (1) the information contained in a model is complete and unchanging whereas the information possessed by a person listening to an utterance is incomplete and may be changed by the understanding of that utterance, and (2) the models are usually presumed to be infinite, whereas a person possesses only finite information. Dynamic model-theoretic semantics [Friedman, Warren, and Moran 1978, 1979; Moran 1980] addresses these problems by allowing the models to contain incomplete information and to have information added to the model. A dynamic model is a "good enough" approximation to an infinite model when it contains the finite subset of information that is needed to determine the meanings of the sentences actually presented to the system.

Dynamic model-theoretic semantics allows the evaluation of a formula to cause the addition of information to the model. This interaction of the evaluation of a formula and the expansion of the model produces several linguistically interesting side-effects, and these have been labelled model-theoretic pragmatics [Moran 1980].

One of these effects occurs when the information given by an element of the model is expanded between the time when that element is identified as the denotation of a sub-expression in the formula and the time when it is used in combination with other elements. If the expansion of the model is not properly managed, the result of the evaluation of such a formula can be wrong (i.e. inconsistent with the contents of the model). Two mechanisms for maintaining the correctness of the denotational relationship are presented. In the first, the management of the relationship is external to the model. This mechanism has the disadvantage that it involves high overhead – the denotational relationships must be repeatedly verified, and unnecessary expansions of the model may be performed. The second mechanism is similar to the first, but eliminates much of this overhead: it incorporates the management of the denotational relationship into the model by augmenting the model's structure.

It is this second mechanism that is of primary interest. It was added to the system to eliminate a source of immediate errors, but it was found to introduce long-term "errors". These errors are interesting because they are the kinds of errors that people frequently make. The structure added to the model permits it to contain inconsistent pieces of information (the structure of a conventional model prevents this), and the mechanism provides a motivated means for controlling which inconsistencies may and may not be entered into the dynamic model.

An important subclass of the inconsistencies provided by this mechanism are known as intensional

substitution failure and this mechanism can be viewed as a variant of both the "impossible" worlds [e.g. Cresswell 1973: 39-41] and the intensional isomorphism [e.g. Lewis 1972] approaches. Since intensionality alone does not provide an account for intensional substitution failure, this mechanism provides an improved account of propositional attitudes.

II. THE PROBLEM

Dynamic models contain incomplete information, and the sets, relations, and functions in these models can be incompletely specified (their domains are usually incomplete). In PTQ, some phrases translate to λ-expressions; other λ-expressions are used to combine and reorder subexpressions. The possible denotations of these λ-expressions are the higher-order elements of the model (sets, relations, and functions). For example, the proper name "John" translates to the logical expression (omitting intensionality for the time being):

(1) $[\lambda P\ P(j)]$

where P ranges over properties of individuals and has as its denotation the set of properties that John has. The sentence "John talks" translates to:

(2) $[\lambda P\ P(j)](talk)$

This formula evaluates to true or false depending on whether or not the property that is the denotation of "talk" is in the set of properties that John has.

The dynamic model that is used to evaluate (2) may not contain the element that is the denotation of "talk". If so, a problem ensues. If the formula is evaluated left-to-right, the set of properties denoted by the λ-expression is identified, followed by the evaluation of "talk". This forces the model to expand to contain the property of talking. The addition of this new property expands the domain of the set of properties denoted by "John", thus forcing the expansion of the characteristic function of that set to specify whether or not talking is to be included. However, because the relationship between the λ-expression for "John" and the set of properties denoted is maintained only during the evaluation of the λ-expression (there is no link from the denotation back to the expression that it denotes), there are no restrictions on how the set is to be expanded. Thus, it is possible to define the property of talking to have John talking and to expand the set previously identified as being denoted by "John" to not include talking, or vice versa. If such an expansion were made, the inconsistency would exist only in the evaluation of that particular formula, and not in the model. Subsequent evaluations of the sentence would recompute the denotation of "John" and get the correct set of properties.

This is not a problem with the direction of evaluation - the argument to which the λ-expression is applied may occur to the left of that λ-expression, for example:

(3) $[\lambda R\ R(talk)](\lambda P\ P(j))$

(note: (3) is equivalent to (2) above).

Finding the argument to which the λ-expression is applied before evaluating the λ-expression is not a viable solution for two reasons. First, some λ-expressions are not applied to arguments, but they have the same problem with their denotations changing as the model expands. Second, having to find the argument to which a λ-expression is applied eliminates one of the system's major advantages, compositionality.

III. THE FIRST MECHANISM - EXTERNAL MANAGEMENT

The mechanism that evaluates a formula with respect to a model has been augmented with a table that contains each λ-expression and the image of its denotation in the current stage of the dynamic model. When the domain of the λ-expression expands, the correct denotational relationship is maintained by expanding the image in the table using the λ-expression, and then finding the corresponding element in the model. If the element in the model that was the denotation of the λ-expression was not expanded in the same way as the image in this table, a new element corresponding to the expanded image is added to the model. This table allows two λ-expressions that initially have the same denotation to have different denotations after the model expands. Since the expansion of elements in the model is undirected, an element that was initially the denotation of a λ-expression may expand into an unused element. The accumulation of unused elements and the repeated comparisons of images in the table to elements in the model frequently imposes a high overhead.

IV. THE SECOND MECHANISM - AUGMENTING THE MODEL

The second mechanism for maintaining the correctness of the denotations of λ-expressions basically involves incorporating the table from the first mechanism into the model. In effect, the λ-expressions become meaningful names for the elements that they denote. These meaningful names are then used to restrict the expansion of the named elements; once an element has been identified as the denotation of a λ-expression, it remains its denotation.*

In the first mechanism, when the domain of two λ-expressions does not contain any of the elements that distinguish them, they will have the same denotation, and when such a distinguishing element is added to the model, the denotations of the two λ-expressions will become different. With meaningful names, this is not possible because the denotational relationship between a λ-expression

----

* Meaningful names are also useful for other purposes, such as generating sentences from the information in the model and for providing procedural - rather than declarative - representations for the information in the model [Moran 1980].

and its denotation in the model is permanent. Since the system cannot anticipate how the model will be expanded, if it is possible to add to the domain of two λ-expressions an element that would distinguish their denotations, those expressions must be treated as having distinct denotations. Thus, all and only the logically-equivalent expressions should be identified as having the same denotation. If two equivalent expressions were not so identified, their denotations would be different elements in the model and this would allow them to be treated differently. For example, if "John and Mary" was not identified to be the same as "Mary and John", it would be possible to have the model contain the inconsistent information that "John and Mary talk" is true and that "Mary and John talk" is false. If two non-equivalent λ-expressions were identified as being equivalent, they would have the same element as their denotation. When an element that would distinguish the denotations of these two expressions was added to the model, the expansion of the element that was serving as both their denotations would be incorrect for one of them and thus introduce an inconsistency.

This need to correctly identify equivalent expressions presents a problem because even within the subset of expressions that are the translations of English phrases in the PTQ fragment, equivalence is undecidable [Warren 1979]. It is this undecidability that is the basis of the introduction of inconsistencies into the model. To be useful in a natural language understanding system, this mechanism needs to have timely determinations of whether or not two expressions are equivalent, and thus it will use techniques (including heuristics) that will produce false answers for some pairs of expressions. It is the collection of techniques that is used that determines which inconsistencies will and will not be admitted into the model.*

V. PROPOSITIONAL ATTITUDES AND INTENSIONAL SUBSTITUTIONAL FAILURE

Intensional substitution failure occurs when one has different beliefs about intensionally-equivalent propositions. For example, all theorems are intensionally-equivalent (each is true in all possible worlds), but it is possible to believe one proposition that is a theorem and not believe another. The techniques used by the second mechanism to identify logically-equivalent formulas can be viewed as similar to Carnap's intensional isomorphism approach in that it is based on finding equivalences between the constituents and the structures of the expressions being compared. This mechanism can also be viewed as using an

----
* While the fragment of English used in PTQ is large enough to demonstrate the introduction of inconsistent information, it is viewed as not being large enough to permit interesting claims about what are useful techniques for testing equivalences. Consequently, this part of the mechanism has not been implemented.

"impossible" worlds approach: if two intensionally-equivalent formulas are not identified as being equivalent, the mechanism "thinks" that it is possible to expand their domain to include a distinguishing element. Since the formulas are equivalent in all possible worlds, the expected distinguishing element must be an "impossible" world.

The presence of intensional substitution failure is one of the important tests of a theory of propositional attitudes. This mechanism is a correlate of that of Thomason [1980], with the addition of meaningful names to intensional objects serving the same purpose as Thomason's additional layer of types.

VI. REFERENCES

Cresswell, M. J. (1973) Logic and Languages, Methuen and Company, London.

Friedman, J., D. Moran, and D. Warren (1978) "An interpretation system for Montague grammar", American Journal for Computational Linguistics, microfiche 74, 23-96.

Friedman, J., D. Moran, and D. Warren (1979) "Dynamic interpretations", Computer Studies in Formal Linguistics report N-16, Dept. of Computer and Communication Sciences, The University of Michigan; earlier version presented to the October 1978 Sloan Foundation Workshop on Formal Semantics at Stanford University.

Lewis, D. (1972) "General semantics", in D. Davidson and G. Harman (eds.) (1972) Semantics of Natural Language, D. Reidel, Dordrecht, 169-218; reprinted in B. H. Partee (ed.) (1976) Montague Grammar, Academic Press, New York, 1-50.

Montague, R. (1973) "The proper treatment of quantification in ordinary English" (PTQ), in J. Hintikka, J. Moravcsik, and P. Suppes (eds.) (1973) Approaches to Natural Language, D. Reidel, Dordrecht, 221-242; reprinted in R. Montague (1974) Formal Philosophy: Selected Papers of Richard Montague, edited and with an introduction by Richmond Thomason, Yale University Press, New Haven, 247-270.

Moran, D. (1980) Model-Theoretic Pragmatics: Dynamic Models and an Application to Presupposition and Implicature, Ph.D. dissertation, Computer Studies in Formal Linguistics, Dept. of Computer and Communication Sciences, The University of Michigan.

Thomason, R. H. (1980) "A model theory for propositional attitudes", Linguistics and Philosophy, 4, 1, 47-70.

Warren, D. (1979) Syntax and Semantics in Parsing: An Application to Montague Grammar, Ph.D. dissertation, Computer Studies in Formal Linguistics report N-18, Dept. of Computer and Communication Sciences, The University of Michigan.