KNOWING INTENSIONAL INDIVIDUALS, AND REASONING ABOUT KNOWING INTENSIONAL INDIVIDUALS

Anthony S. Maida

Center for Cognitive Science
Box 1911, Brown University
Providence, Rhode Island 02912, USA

ABSTRACT

This paper outlines an approach toward computationally investigating the processes involved in reasoning about the knowledge states of other cognitive agents. The approach is Fregean and is compared with the work of McCarthy and Creary. We describe how the formalism represents the knowing of intensional individuals, coreferentiality, iterated propositional attitudes, and we describe plans to test, the scheme in the domain of speech act recognition.

I INTRODUCTION

Humans quite effectively reason about other humans' knowledge states, belief states, and states of wanting. Unfortunately, the processes by which humans do this are not well understood. This paper outlines an approach toward computationally investigating these processes. This approach involves two components, the first of which involves adequately representing knowledge about others' knowledge; and the second of which involves describing implementable processes by which it is possible to reason about such knowledge. Our approach is Fregean to the extent that the kind of cognitive system we propose puts emphasis upon the representation of Fregean senses. However, the approach is not entire]y Fregean because we do not represent denotations. This contrasts with the purely Fregean approaches of McCarthy (1979) and Creary (1979).

A. McCarthy's Approach

McCarthy begins with the simple example of Pat knowing Mike's phone number which is Incidentally the same as Mary's phone number, although Pat does not necessarily know this. This example immediately exposes one of the difficulties of reasoning about knowledge, namely, the problem of inhibiting substitution of equal terms for equal terms in referentially opaque contexts. McCarthy's approach toward solving this problem involves explicitly representing senses and denotations.

B. Creary's Extension

Creary extended McCarthy's system to handle iterated propositional attitudes. McCarthy's system fails for iterated propositional attitudes because propositions are represented but not their concepts. Creary's extensions involve introducing a hierarchy of typed concepts. Thus for individuals such as the person Mike, this scheme would have the person Mike, the concept of Mike, the concept of the concept Mike, and so forth. The higher concept is the Fregean sense of the lower concept, which reciprocally is the denotation of the higher concept. A similar situation holds for propositions. The hierarchy would consist of a truth value, the proposition which denotes the truth value, the concept of that proposition, and so on. This scheme allows for the representation of iterated propositional attitudes because all objects in the domain of discourse (most notablv propositions) have senses.

C. The Maida-Shapiro Position

Our starting point is the observation that knowledge representations are meant to be part of the conceptual structure of a cognitive agent, and therefore should not contain denotations. The thread of this argument goes as follows: A cognitive agent does not have direct access to the world, but only to his representations of the world. For instance, when a person perceives a physical object such as a tree, he is really apprehending his representation of the tree. Hence, a knowledge representation that is meant to be a component of a "mind" should not contain denotations. A more elaborate statement of this position can be found in Maida and Shapiro (1982) and the system for representing knowledge, *called* Lambda Net, described in the remainder of this paper is described in Maida (1982). For our purposes, refraining from representing denotations achieves two goals: 1) the problem of substitution of equal terms for equal terms goes away because distinct terms are never equal; and 2) we can represent iterated propositional attitudes without invoking a hierarchy of types.

II LAMBDA NET

A. Intensional Individuals

There is a class of intensional individuals for which it can be said that they have a value as seen in assertions such as:

a) John-bear knows where Irving-bee is.
b) John knows Mike's phone number.
c) John knows the mayor's name.

What does John know in each of these sentences? He knows the value of some intensional individual. We can characterize these individuals by observing that they each involve a two-argument relation; namely, location-of, phone-no-of, and name-of, respectively. In each case, one argument is specified; namely: Irving-bee, Mike, and the mayor. The other argument is unspecified. We make the assumption that context uniquely determines the value of the unspecified argument. This value is the value of the intensional expression. The expressions themselves can now be represented as:

```
d) (the (lambda (x) (location-of Irving x)))
e) (the (lambda (x) (phone-no-of Mike   x)))
f) (the (lambda (x) (name-of     mayor  x)))
```

B. Knowing Intensional Individuals

Since each of these expressions has a value, someone can know their values. We will express this via a relation called "know-value-of" which takes a cognitive agent and an intensional individual as arguments. To represent "John knows Mike's phone number," we write:

```
g) (know-value-of John
       (the (lambda (x) (phone-no-of Mike x))))
```

Observe that we treat propositional attitudes, and attitudes toward intensional individuals, as being relational and not as intensional operators. Knowing is viewed as correct (but not necessarily justified) belief.

The meaning of "know-value-of" entails that if John knows the value of Mike's phone number, and the value of Mike's phone number is 831-1234, then John "knows-that" the value of Mike's phone number is 831-1234.

C. Iterated Propositional Attitudes

Reasoning about the knowledge states of others necessarily involves iterated propositional attitudes because the cognitive agent doing the reasoning is generating beliefs about another agent's knowledge state which itself may contain beliefs about the beliefs of other cognitive agents. Thus it is useful to show how Lambda Net represents such assertions. Creary (1979) offers three semantic interpretations of the ambiguous sentence:

h) Pat believes that Mike wants to meet Jim's wife.

He suggests that the task of representing these interpretations provides a strong test of the representation. In order to allow the reader to compare the Lambda Net scheme with Creary's we list the representations below. In each case, we give a rendering of the interpretation in English, our representation, and Creary's representation.

1) Pat believes that Mike wants to meet Jim's wife as such.

```
a) (believe-that Pat
       (wants Mike
          (meet
             (the (lambda (x:person)
                     (wife-of Jim x))))))

b) believes (pat, Wants{Mike, Meet$
       {Mike$, Wife$ Jim$}})
```

2) Pat believes that the person Mike wants to meet is Jim's wife, although Mike doesn't necessarily believe that.

```
a) (believe Pat
       (wife-of Jim
         (the (lambda (x:person)
           (wants Mike (meet Mike x))))))

b) believes (pat, Exist P$.Wants {Mike,
   Meet${Mike$, P$}} And Conceptof
   {P$, Wife Jim})
```

3) There is a specific person Pat believes Mike wants to meet. Neither necessarily believes this person is Jim's wife, but it incidentally is.

```
a) (wife-of Jim
       (the (lambda (x:person)
          (believe Pat (want Mike
                         (meet Mike x))))))

b) ∃P$ P.believes(pat, Wants{Mike,
   Meet${Mike$, P$}}) & conceptof(P$,P) &
   conceptof(P, wife jim)
```

The reader should refer to the original papers, Creary (1979) and Maida (1982), to make the proper comparison. One of Creary's goals is to stay within the confines of a first-order logic. Lambda Net does not have that constraint.

D. Knowing Coreferential Intensional Individuals

To assert that two intensional individuals are coreferent, we write:

```
i) (equiv individual-1 Individual-2)
```

The relation "equiv" is mnemonic for extensional equivalence, and is the only reference to extensionality used in Lambda Net. One of our performance goals is to design a system which reacts appropriately to assertions of coreference. This involves specifying a method -to treat transparent and opaque relations appropriately. A relation, or verb, such as "dail" or "value-of" is transparent whereas a relation such as "know" is opaque with respect to its complement position. We can express this as:

```
(transparent dial)
(transparent value-of)
(conditionally-transparent know lst-arg 2nd-arg)
```

"Dial" and "value-of" are unequivically transparent, whereas "know" (either know-that or know-value-of) is transparent on the condition that the agent doing the knowing also knows that two entities are coreferent. We can partially express

E. Axiom of Rationality

A system that reasons about the beliefs of another cognitive agent must make assumptions about the rationality of that agent in regard to what he considers legitimate rules of inference. We shall assume that all cognitive agents utilize the same set of inference schema. This is the Axiom of Rationality and we further assume that this set of schema is exactly the set given in this paper. A statement of the Axiom of Rationality is:

> Axiom of Rationality - If a cognitive agent knows or is capable of deducing all of the premises of a valid inference, then he is capable of deducing the conclusion of that inference.

The Axiom of Rationality enables one cognitive agent to determine by indirect simulation whether another cognitive agent is capable of inferring something. It implies, "If I figured it out and he knows what 1 know, then he can also figure it out if he thinks long enough." We will assume that the situations involved in knowing about telephone numbers are simple enough to make plausible the stronger rule, "If 1 figured out and he knows what I know, then he has definitely figured it out."

F. Reasoning about Knowing

In this section we give an example of how reasoning about knowing can take place in Lambda Net by modeling the following situation involving a propositional attitude.

Premises: 1) John knows that Pat knows Mike's phone number.
2) John knows that Pat knows that Mike's phone number is the same as Mary's phone number.

Conclusion: John knows that Pat knows Mary's phone number.

By the definition of knowing as correct belief, it follows that: 1) Pat knows Mike's phone number; and, 2) Pat knows that Mike's phone number is the same as Mary's phone number. From conditional transparency and the Axiom of Rationality, the conclusion follows.

III SUMMING UP

A. What has been Achieved?

A system which can reason validly about knowledge must have at least the following three performance characteristics: 1) The system must be able to represent assertions involving iterated propositional attitudes and reason from these assertions; 2) The system must react appropriately to assertions involving coreference between distinct intensional individuals; and, 3) The system must felicitously represent that another cognitive agent can know the value of some intensional individual without the system itself necessarily knowing the value. Lambda Net has these characteristics just as Creary's (1979) does. However, Lambda Net offers the advantage of not invoking a hierarchy of conceptual types in order to achieve these performance characteristics.

B. Current Work

We are implementing this system to process speech acts using the general strategy described by Allen (1979). This approach views speech acts as communications between cognitive agents about obstacles and potential solutions to achieving some goal. Therefore, comprehending and appropriately reacting to a speech act necessarily requires the capacity to reason about another cognitive agent's goals (wants), planning strategy, and knowledge states.

REFERENCES

1 Allen, J. A plan-based approach to speech act recognition. Ph.D. Thesis, Computer Science, University of Toronto, 1979.

2 Creary, L. "Prepositional attitudes: Fregean representation and simulative reasoning." In Proc. IJCAI-79. Tokyo, Japan, August, 1979, pp."176-181.

3 Maida, A. "Using lambda abstraction to encode structural information in semantic networks." Report //1982-9-1, Box 1911, Center for Cognitive Science, Brown University, Providence, Rhode Island, 02912, U.S.A.

4. Maida, A. and Shapiro, S. "Intensional concepts in propositional semantic networks." Cognitive Science 6:4 (1982) 291-330.

5 McCarthy, J. "First order theories of individual concepts and propositions." In J. Hayes & D. Michie (Eds.) Machine Intelligence 9, New York: Halsted Press, 1979.