AD-A102 765 BOLT BERANEK AND NEWMAN INC CAMBRIDGE MA F/G 5/2

IT'S FOR YOUR OWN GOOD: A NOTE ON INACCURATE REFERENCE.(U)

JUL 81 C R PERRAULT, P R COHEN N00014-77-C-0378

UNCLASSIFIED BBN-4723 NL

1 OF 1

AD A102765

END

DATE FILMED

9-81

DTIC

AD A102765

DTIC FILE COPY

Bolt Beranek and Newman Inc.

LEVEL II

(12)

Report No. 4723
Technical Report

It's for Your Own Good: A Note on Inaccurate Reference

C.R. Perrault and P.R. Cohen

DTIC
ELECTE
AUG 12 1981
S D
C

July 1981

Submitted to:
Advanced Research Projects Agency
Department of Defense

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

81 8 12 044

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report	2. GOVT ACCESSION NO. AD-A102	3. RECIPIENT'S CATALOG NUMBER 165
4. TITLE (and Subtitle) It's For Your Own Good: A Note on Inaccurate Reference		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER BBN Report No. 4723
7. AUTHOR(s) C. R. Perrault & P. R. Cohen		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0378
9. PERFORMING ORGANIZATION NAME AND ADDRESS Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Massachusetts 02238		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE July 1981
		13. NUMBER OF PAGES 30 pages
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited.
It may be released to the Clearinghouse, Department of Commerce for sale to the general public.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Reference, Mutual Belief, Planning, Speech Acts

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In his book Speech Acts, Searle suggests that his analysis of illocutionary acts can be extended to accounts for the "propositional act" of reference. He wants to give necessary conditions for a speaker successfully referring to some entity by using a certain referring expression in an utterance to a certain hearer. We point out here some inadequacies in

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Searle's conditions, in particular how they fail to account for cases of successful reference through expressions which speaker (and hearer) may believe are not true of their intended referent. More specific conditions based on the notion of mutual belief are proposed.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist
A

Report No. 4723 Bolt Beranek and Newman Inc.

It's For Your Own Good: A Note on Inaccurate Reference

C. Raymond Perrault
Dept. of Computer Science
University of Toronto
Toronto Canada

and

Philip R. Cohen
Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

This research was supported in part by the National Research Council of Canada under Operating Grant A9285, and by the Advanced Research Project Agency of the Department of Defense and was monitored by the Office of Naval Research under Contract N00014-77-C-0378. This paper has appeared in Elements of Discourse Understanding, A. Joshi, B. Webber, and I. Sag (Eds.), Cambridge: Cambridge University Press, 1981.

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	3
2.	BELIEF	9
3.	THE NEED FOR MUTUAL BELIEF	11
4.	OVERRIDING MUTUAL BELIEF	13
5.	A RECONSIDERATION	19
6.	THE ATTRIBUTIVE CASE	23
7.	CONCLUDING REMARKS	27

ABSTRACT

In his book Speech Acts, Searle suggests that his analysis of illocutionary acts can be extended to account for the "propositional act" of reference. He wants to give necessary conditions for a speaker successfully referring to some entity by using a certain referring expression in an utterance to a certain hearer. We point out here some inadequacies in Searle's conditions, in particular how they fail to account for cases of successful reference through expressions which speaker (and hearer) may believe are not true of their intended referent. More specific conditions based on the notion of mutual belief are proposed.

ACKNOWLEDGEMENTS

Many thanks to Alex Borgida, David Israel, Corot Reason, Candy Sidner, and Bonnie Webber for their comments, and to Norma Peterson for assistance in preparing the paper.

1. INTRODUCTION

Austin [1] was one of the first to emphasize the distinction between the truth value of a proposition and the use of that proposition within an utterance that is the result of a (speech) act performed by some speaker for some hearer(s). Propositions can be true or false and the study of their relation to the world is the domain of classical semantics. Acts can succeed or fail. Their success may depend on certain circumstances obtaining, in particular on the speaker holding certain beliefs and having certain intentions. Grice [7] gives an account of what a speaker means when performing an act of communication in terms of the speaker's intention that the hearer should recognize certain intentions of the speaker. Strawson [18] and Searle [16] propose slightly different ways of applying Grice's theory to define the illocutionary acts first discussed by Austin.

Following his account of illocutionary acts, Searle also suggests that the sentence-meaning/speaker-meaning distinction can be extended to reference, i.e. that there is a difference between what a definite description, say, refers to, and what a speaker intends to refer to by using that description. He then proposes conditions defining the felicitous performance of the reference act.

Cohen and Perrault [3] and Perrault and Allen [13] show how certain difficulties with Searle's definitions of the illocutionary acts REQUEST and INFORM can be overcome by redefining them as operators in a problem-solving system (e.g. Fikes & Nilsson [6]). Cohen's [4] OSCAR program can construct sequences of actions by which one agent can achieve certain goals, and these sequences can involve the performance of

REQUESTs and INFORMs. While OSCAR constructs the propositional content of its illocutionary acts (i.e. what act is being requested, or of what proposition the hearer is to be informed), it does not construct noun phrases by which the speaker can refer to entities as part of performing the referring act.

This leads us to seek necessary and sufficient conditions defining when a speaker S can be said to have referred to an entity x in uttering a referring expression E. Much of the difficulty stems from the fact that although the classic examples of referring expressions are proper names and definite descriptions, not all utterances of expressions of these grammatical types are normally said to refer. Well known examples to the contrary are "Cerebus" and "the Golden Mountain" which presumably never referred, and "the largest prime number" which could not. Some definite descriptions such as "the Prime Minister" can be used to refer to an individual, as in "I met the Prime Minister yesterday", but can also be used intentionally, as in "The Prime Minister is the head of the executive". In the latter use, the truth conditions of the sentence are independent of what individual the definite description identifies, or even of whether there is such an individual.

Searle's analysis, like that of Russell [14] and Strawson [18] before him, is limited to singular definite reference whereby the speaker is assumed to be trying to identify some existing entity for the hearer. We adopt here a similarly restricted view.

It is convenient to accept Searle's distinction between "fully consummated reference", one in which the identification of an object is communicated successfully to the hearer, and "successful reference" where the speaker had all the right

intentions and conformed to all the right conventions, although, to use Austin's term, uptake may not have been secured. Restricting himself to the latter, Searle states his "principle of identification" (PI):

> (PI.1) A necessary condition for the successful performance of a definite reference in the utterance of an expression is that either the expression must be an identifying description or the speaker must be able to produce an identifying description on demand. (Searle [16], page 88)

In this paper we claim that Searle's PI fails to take into account the fact that <u>what</u> the description can be used to identify depends on the beliefs of the speaker and hearer, including the speaker's beliefs about the hearer, etc. In particular it does not account for cases where the description used in a successful reference may not only not be satisfied by its intended referent but may be believed by the speaker not to be satisfied. In the rest of this section we consider a few preliminary objections to PI.1. Section 2 contains a brief description of the properties of the propositional attitude "belief" on which the rest of the analysis relies. Section 3 reviews some arguments of Clark and Marshall [2] who claim that an <u>infinite</u> set of beliefs about the entity's satisfying the identifying description is necessary. In section 4 we argue that Clark and Marshall's claim is too strong by giving a series of counterexamples. The section concludes with a refined version of the PI. Section 5 suggests that <u>how</u> the description identifies the intended referent must also be considered. Section 6 argues that the version of the PI given in section 5 does not apply to the so-called "attributive" uses of descriptions, and section 7 concludes the paper.

Preliminary Objections

Returning to PI.1, Searle's analysis of what counts as an "identifying" description is rather vague. He claims that for a speaker to have identified an object by means of a description means that "there should be no longer any doubt or ambiguity about what exactly is being talked about ... questions like 'who?', 'what?' and 'which one?' are answered." These answers can be provided in two ways: by demonstrative presentation and by "descriptions in purely general terms which are true of the object uniquely". Descriptions may also rely on a mixture of demonstrative devices and descriptive predicates. "So identification rests squarely on the speaker's ability to supply an expression of one of these kinds, which is satisfied uniquely by the object to which he intends to refer. I shall hereafter call any such description an identifying description". (emphasis in the original).

Nowhere does Searle discuss what he means by a description being "satisfied uniquely" by an object. One object of this paper is to explore this question.

First, it is clear that "satisfying uniquely" must be considered with respect to some "context" or "focus" (as for example described by Grosz [8]), created by the conversational process, the physical setting, and probably cultural conventions. The referent of "the man" in "John met a man in the street. He gave the man a dime." is the man whom John met and who was mentioned in the previous sentence (Webber [19]). Searle's claim in his principle of identification that the speaker should provide an identifying description or "be able to produce one on demand" is meant to capture the fact that the speaker should be able to expand the description by explicitly including as much of the context as necessary to identify the object.[1]

---

[1]One of Searle's objectives is to show that the principle of identification is a special case of his earlier "principle of expressibility" which stated roughly is "what can be meant can be said."

Second, it is also clear that the use of a referring expression by a speaker has little or nothing to do with whether it is true that the object he intends to refer to actually satisfies the description and does so uniquely. We can all refer to Roger Bannister as "the first man to run a mile in less than four minutes" even if there may have been a Bantu tribesman who had a long way to go to the nearest tree in escaping one of the neighborhood beasts.

With these refinements in mind, we can now reformulate the principle of identification as follows:

> (PI.2) A necessary condition for the successful performance of a definite reference by a speaker S using a description D in a context C is that S believes that D is fulfilled in C.

The introduction of the context C is intended here to account for Searle's "or must be able to produce an identifying description on demand". What context is remains a problem: we will take it to be a set of entities "known" to speaker and hearer. By "D is fulfilled in C" we mean that exactly one entity in C satisfies D, i.e. (Ex) (Ay:C) D(y) => x = y.

In the rest of this paper we want to show that this version of the principle of identification is neither sufficient nor necessary. It is not sufficient because it says nothing about the knowledge which H must <u>share</u> with S about D. This point is discussed by Clark and Marshall [2], and their evidence will be reviewed below. It is not necessary because in some sense, to be made precise later, it is possible for the speaker to believe (and for the speaker to believe the hearer believes ...) that D is <u>not</u> fulfilled in C, or that it is fulfilled by the "wrong" entity and yet still use the description in an essential way.

2. BELIEF

It will be convenient to formulate the Principle of Identification in terms of statements in a logic of belief. Following Hintikka [9] we interpret belief as a modal operator B(a,P) where a is the believing agent, and P the believed proposition. This is usually written aB(P) or aBP and satisfies the following axiom schemas, where a and P range over agents and propositions, respectively:

B.1 aB(all axioms of the propositional calculus)

B.2 aB(P) => aB(aB(P))

B.3 aB(P) => <u>not</u> aB(<u>not</u> P)

B.4 aB(P => Q) => (aB(P) => aB(Q))

B.5 Ex aB(P(x)) => aB(Ex P(x))

B.6 all agents believe that all agents believe B.1 to B.5

It is important to note that the converse of B.5 does not hold.

A sound and complete model for these axioms may be given in terms of possible worlds as suggested by Hintikka [9].

3. THE NEED FOR MUTUAL BELIEF

Clark and Marshall [2] give a series of examples[2] which show that for S to refer for H to some entity E using some description D in a context C it is not sufficient that S believe that D is fulfilled in C, but S must also believe that H believes it does, and that H believes S believes it does, etc. For S to ask H "How did you like the movie?" it is not sufficient for S to believe that H went to exactly one movie; S must also believe that H believes H went to exactly one movie, etc. If any of these conditions fail before the reference act is made, then in order to successfully refer the speaker must be willing to accept that the hearer will attribute all of them to her. Thus no finite conjunction of the form

SB(P) & SBHB(P) & ... & SBHB...SB(P)

is sufficient for successful reference, where P is the proposition "E fulfills D in C". We will say that S and H mutually believe that P (written MB(S,H,P)), if the infinite conjunction

(1) SB(P) & SBHB(P) & SBHBSB(P) & ...

obtains.[3] We will show in the next section that although an

---

[2]A similar argument is made by Schiffer [15] to show the lack of sufficiency of Strawson's [18] refinement of Grice's [7] analysis of speaker meaning.

[3]Notice that MB(H,S,P) interchanges H and S in (1) and thus MB(S,H,P) and MB(H,S,P) are not equivalent. However MB(S,H,P) is equivalent to

SB(P) & SB(MB(H,S,P)).

infinite number of conjuncts <u>are</u> necessary for successful reference, mutual belief that the description be fulfilled is not. But first, we digress briefly to discuss mutual belief.

Mutual belief is a way of representing facts that humans acquire because they expect other humans' perceptions and deductive processes to be similar to theirs. Lewis [10] and Schiffer [15] noted that when two agents S and H together witness some event A (or, to use Clark and Marshall's term, are <u>copresent</u> at A) then an unbounded set of propositions seems to be acquired by both. From S's point of view,

S believes that A occurred,

and since S saw H witnessing A,

S believes H believes A occurred,

and since S saw H see S witnessing A,

S believes H believes S believes A occurred,

and so forth. Assuming rationality on the part of S and H, S must agree to all these propositions, i.e. S and H mutually believe that A occurred.

As a consequence of their copresence and their assumption of mutual rationality, S and H may also be assumed to acquire consistent descriptions of entities involved in some event. For example, if S and H together see a table on which sit a green block G and a red block R then it is reasonable to conclude

MB(S,H,R fulfills "the red block on the table") and

MB(S,H,G fulfills "the green block on the table").

4. OVERRIDING MUTUAL BELIEF

Consider now the following example adapted from Donnellan [5]:

> Example 1: S and H are at a party. They watch together as water and gin are being poured in two identical glasses and given to women W1 and W2 respectively. Unbeknownst to H, S sees W1 and W2 exchange glasses. Later S tells H "The woman with the martini is the mayor's daughter."

There is no doubt that in so doing S successfully referred to W2 even though W2 was not drinking a martini, nor did S believe she was, although S believed H believed she was. S could not have been referring to W1 because S does not believe that H could recognize the woman with the martini as referring to W1, since S's knowledge of the exchange is not shared with H. In some sense, S's utterance is misleading since a (perlocutionary) effect of the assertion is that H believes that S believes that W2 is the mayor's daughter. However, neither S's nor H's beliefs about who the woman with the martini is need change as a consequence of S's securing uptake for the referent of "the woman with the martini". S's only fault is in not correcting a previous misunderstanding.

S is thus relying on S and H having shared the drink pouring experience to construct descriptions of W1 and W2 they could

agree to.[4] Immediately following the pouring of the drinks it is true that

MB(S,H,W1 fulfills "TWWW" & W2 fulfills "TWWM").[5]

Beliefs about objects and their descriptions can also be acquired privately and may override some of the conjuncts of the mutual belief which results from two agents witnessing an event together. In Example 1, S saw W1 and W2 exchange glasses but S does not believe H saw the exchange. Thus the information available to S immediately prior to her assertion can be represented as:

---

[4]Searle recognizes that a reference may be successful without the description being true of the object referred to and quotes an example of Whitehead [20]:

S:"That criminal is your friend"

H:"He is my friend and you are insulting"

He dismisses it because "the word 'that' in 'that criminal' indicates the object either is present or has already been referred to by some other referring expression and that the present reference is parasitic on the earlier. The descriptor 'criminal' is not essential to the identification ... which is achieved by other means." (p.90)

In Example 1 however, "the woman with the martini" is not parasitic for although S, H and W are copresent, no use is made, nor need be made, of deictic expressions for successful reference. The example would work equally well if S and H were each watching the group including W from separate rooms on TV monitors and communicating by telephone.

[5]For the rest of the paper, "TWWM" will be used as an abbreviation for "the woman with the martini" and "TWWW" will be used as an abbreviation for "the woman with the water".

SB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SB(MB(H,S,W1 fulfills "TWWW" & W2 fulfills "TWWM")).

One might claim that S's reference to W2 was successful because SBHB(W2 fulfills "TWWM"), since this follows from SB(MB(W2 fulfills "TWWM")). However, this condition is still too strong. In the rest of this section, we give a series of examples which show that any finite number of conjuncts of the formula

MB(S,H,W1 fulfills TWWW & W2 fulfills TWWM)

can fail and yet S can still refer to W2 as TWWM.

> Example 2: S and H are at a party. They watch together as water and gin are being poured in two identical glasses and given to women W1 and W2 respectively. Later S sees H see the women swap glasses but S believes that H did not see S see H. S then tells H "TWWM is the mayor's daughter".

We claim that S has successfully referred to W2 and that before she made her assertion it was the case that:

SB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBMB(S,H,W1 fulfills "TWWW" & W2 fulfills "TWWM").

As in Example 1, S could not be referring to W1 since to do so would require S to expect H to understand the reference based on H's private beliefs. But following an argument similar to that given after Example 1, H cannot use private beliefs to understand a reference which S is trying to make. Thus the description may fail in SB and in SBHB. In the next example, it also fails in SBHBSB.

> Example 3: S and H are at a party. They watch together as water and gin are being poured in two identical glasses and given to women W1 and W2 respectively. Later S sees H see the women swap glasses, without seeing H see

him. S also overhears A telling H that S saw him see the exchange. Later, S tells H:"TWWM is the mayor's daughter."

Again here S has made a successful reference to W2 and before the assertion it also was the case that:

SB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBSB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBSBMB(H,S,W1 fulfills "TWWW" & W2 fulfills "TWWM").

By now it should be obvious that this game can be played forever. True believers can skip Example 4.

Example 4: S and H are at a party. They watch together as water and gin are being poured in identical glasses and given to W1 and W2 respectively. Later S sees H seeing the women exchange glasses, but S believes H did not see her see the exchange. A, whom S believes to be truthful, tells S that A told H that S saw H see the exchange. S knows that H is listening to their conversation but pretends not to notice. S then tells H:"TWWM is the mayor's daughter."

Here we have:

SB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBSB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBSBHB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBSBHBMB(W1 fulfills "TWWW" & W2 fulfills "TWWM")

In Examples 1-4, reference has been successful, although at some of the "early" levels SB, SBHB, etc. the desired referent failed to fulfill the description. Beyond this finite set of conjuncts however, there is mutual belief that the referent does

fulfill the description. Can a reference act succeed if the description fails at an unbounded number of levels? We claim not, although intuitions differ on this point.

> Example 5: Two women W1 and W2 are holding martini glasses. S thinks that W1's glass contains water and W2's contains gin, and she has told this to H who has replied that he believes it to be the other way around. Neither is convinced by the other's argument, and neither thinks one would lie to the other. S tells H:"The woman with the martini is the mayor's daughter".

Here we claim that S has failed to refer to either W1 or W2, and that S's situation is as follows:

SB(W1 fulfills "TWWW" & W2 fulfills "TWWM") &

SBHB(W1 fulfills "TWWM" & W2 fulfills "TWWW") &

SBHBSB(W1 fulfills "TWWW" & W2 fulfills "TWWM") &

....

Since SB(P) => SB(SB(P)), the infinite conjunction above implies

MB(S,H,SB(W1 fulfills "TWWW" & W2 fulfills "TWWM")) &

MB(S,H,HB(W1 fulfills "TWWM" & W2 fulfills "TWWW")).

Note that S could refer to W2 by asserting (11) for then the description is again a mutually believed one.

"The woman that I believe is holding a martini is the mayor's daughter."

Let us say that S and H agree that P (written AGREE(S,H,P)) if

MB(S,H,P) or SBMB(H,S,P) or SBHBMB(S,H,P) ...

Notice that AGREE(S,H,P) implies SB(AGREE(H,S,P)).

Examples 1-5 can be accounted for by the following version

of the principle of identification:

> (PI.3) A necessary condition for a speaker S to refer for H to some entity E using a description D in a context C is that AGREE(S,H, E fulfills D in C).

Before re-examining the martini examples, note that in them all beliefs are acquired visually. But clearly mutual belief can be acquired verbally as well, and the same "overriding" phenomena can occur. For example, if S tells H "That is the woman with the martini" pointing at E, then if H and S mutually believe S to be sincere and correct, we have MB(H,S,E fulfills TWWM). If H believes S to be sincere but wrong then we have

HB(E does not fulfill TWWM) &

HBMB(S,H,E fulfills TWWM)

If H believes S to be lying then

HB(E does not fulfill TWWM) &

HBSB(E does not fulfill TWWM) &

HBSBMB(H,S,E fulfills TWWM).

In any case, E can subsequently be referred to as "the woman with the martini."

5. A RECONSIDERATION

In our interpretation of the martini examples so far we have been assuming that the speaker intends the description "the woman with the martini" to be interpreted with respect to the information S and H share about the event during which the glasses were filled, and with respect to general (shared) information such as "if a person is holding a glass containing a martini at some time t in the past, and if the person continues to hold that glass until the present, and if no change is made to the contents of the glass in the interval, then the person is holding a martini now". Let us call this proposition J the <u>justification</u> for the reference.

In the cocktail party context, S clearly does not intend that the referent of "the woman with the martini" be determined by H's actually <u>testing</u> the contents of the glasses. If this were S's intention, then S could not refer to the woman who is actually holding the water as "the woman with the martini". Thus S's ability to use descriptions which are inaccurate but which he knows to be inaccurate depends crucially on <u>how</u> the referent is to be determined by H, or of what justification there is for a referent to fulfill a description. This is not captured by the current version of the principle of identification.

How then are we to relate the entity E that S intends to refer to, the description "the x such that D(x)" uttered by S to H, and the way S intends H to determine the referent? We suggest the following:

> (PI.4) For a speaker S to refer to an entity E by uttering "the x such that D(x)" to H in a context C it is necessary that there exist formulas ID(x) and J(x), called the <u>initial description</u> and <u>justification</u> formulas respectively such that

1. MB(S,H,ID(E) & (Ax:C) ID(x) => x=E)
2. MB(S,H,(Ax)(ID(x) & J(x) => D(x))
3. AGREE(S,H,J(E))
4. AGREE(S,H, (Ax,y in C) D(x) & D(y) => x=y)
5. S intends H to identify E by means of ID and J,
6. S intends H to recognize S's intention 5.

The initial description ID is a predicate which S and H mutually believe is fulfilled by E (Condition 1). The justification predicate J ensures that if an entity satisfies ID and J then it also satisfies D, and this fact is mutually believed (Condition 2). In the martini examples, we take

ID(x) = woman(x) & hold(x,G,T) & martini(G,T),

J(x) = (At:[T,NOW]) hold(x,G,t) & martini(G,t)

and

D(x) = (Eg:glass) & woman(x) & hold(x,g,NOW)

& martini(g,NOW),

where hold(s,g,t) is true if x holds glass g at time t, and martini(g,t) if glass g contains a martini at time t. We take G, T, and NOW to be constants denoting the glass E held, the time at which the martini was poured, and the current time respectively. [T,NOW] denotes the interval between T and NOW. The truth of condition 2 then follows as a special case of the "frame axiom"

(Ax) (Ag) (At1) (At2) hold(x,g,t1) & martini(g,t1) &

(At:[t1,t2]) hold(x,g,t) & martini(g,t)

=> hold(x,g,t2) & martini(g,t2)

which is mutually believed by everyone.

Condition 3 states that S and H AGREE that the justification holds of E; in our case this means AGREEing that E held a martini between the time of pouring and the present. This condition is

not mutually believed. Condition 4 states that S and H AGREE that at most one entity satisfies the uttered description. This condition does not follow from 1, 2, and 3, since for example, another woman with a martini might have walked into the room between times T and NOW.

Conditions 5 and 6 are analogous to the Gricean conditions in Searle's definitions of illocutionary acts.[6] It is not sufficient that ID and J should exist: S must also intend that they be used by H. In a Searle-type definition of an illocutionary act, such a condition also includes the qualification that the recognition of intention be done at least in part because of the utterance itself. It is much more difficult to see how the description itself can suggest to the hearer the ID and J that the speaker intends him to use.

---

[6]Condition 6 is probably not strong to enough to avoid analogues to Schiffer's counterexamples.

6. THE ATTRIBUTIVE CASE

In the examples we've examined so far, the speaker and hearer could both identify the entity being referred to, in the sense that both were assumed to have about the referent information that could not be deduced from the description uttered. Their "acquaintance" with the referent gives them descriptions relating appearance, location, time, etc. Definite descriptions are, however, often used when one or the other party does not have other information to identify the referent, i.e. cannot establish coreference of the definite description uttered with any other one which does not follow from it logically.

For example, a passenger asking an information clerk "What is the departure time of the next train to Montreal?" may not know anything about the referent of "the departure time ..." other than what he has said. He presupposes that a departure time exists, and that it is unique but the uniqueness here is a consequence of the meaning of "departure time". In no way could he identify the referent; in fact, the more he could, the less likely he would be to ask the question. In this case there seems no way that the speaker could inaccurately refer to "the departure time ..." as was the case in the martini examples.

Unfortunately, PI.4 is too strong to accommodate this use of the definite description and we see no way of weakening it while still accounting for the martini examples. Even if ID is taken to be D (and J to be trivially true), condition 1 of PI.4 fails in general since no copresence situation can be assumed to have asserted it previously. At best, MB(S,H,(Ex) (Ay:C) D(y) => x=y) is necessary. The condition AGREE(S,H,(Ex) (Ay:C) D(y) => x=y) seems too weak since SB((Ex) (Ay:C) D(y) => x=y) must be true, and we can find no case where a formula of the form SBHBSB...SB((Ex) (Ay:C) D(y) => x=y) fails.

Thus there are two quite different uses of definite descriptions, as defined by the knowledge conditions necessary to their utterance, and they appear to correspond to Donnellan's "referential" and "attributive" uses. Only if a description is used referentially can the speaker not believe that it is true of the entity he wishes to refer to. (Ironical utterances, as usual, must be treated specially).

The satisfaction of the conditions for referential use do not imply satisfaction of those for attributive use, unless condition 1 happens to be mutually believed, as well as agreed. This would justify defining two different reference acts.

Ortony and Anderson [12] make the very suggestive claim that proper names and definite descriptions each have a different primary role with respect to reference: uses of proper names are primarily referential and uses of definite descriptions are primarily attributive. However, each can indirectly play the role of the other.

This proposal is particularly appealing because it is very close in spirit to several proposals ([17], [11], [13]) for explaining the relation between the literal, or direct interpretation of utterances such as "Can you reach the salt?" (as a yes/no question) and their indirect interpretations (e.g. as a request to pass the salt). Unfortunately the steps by which the illocutionary forms can be related (c.f. Perrault and Allen [13]) do not seem to be the same as those required in the case of reference. However, the prospect of such an explanation is highly enticing.

Defining the reference acts requires specifying the effects of these acts, which traditionally has meant specifying the

weakest formula which becomes true when the act is successfully executed. This requires investigating the relation between the reference acts and the illocutionary acts, and remains to be done.

7. CONCLUDING REMARKS

It is beyond the scope of this paper to discuss the logical, psychological, or computational aspects of mutual belief. The difficulties in giving adequate semantics for belief are well known. We simply want to point out that the countable number of conjuncts in the definition of mutual belief need not make things worse.

If one considers the change in the state of some agent witnessing some event in the presence of another agent, it is reasonable to assume that the mutual belief will be acquired as a "unit" rather than one conjunct at a time, and that certain deductions can also be made, on the basis of mutually believed information, which result in new mutually believed propositions. This can be represented informally as

MB(S,H,P) & MB(S,H,P -> Q) -> MB(S,H,Q).

Clark and Marshall [2] discuss the acquisition of mutual belief and Cohen [4] examines some related computational questions such as data structures which allow finite representations of belief and mutual belief in a program which generates speech acts.

REFERENCES

1. Austin, J. L. How to do things with words. Oxford University Press, London, 1962.

2. Clark, H. H., & Marshall, C. Definite reference and mutual knowledge. In Elements of Discourse Understanding, Joshi, A. K., Sag, I. A., & Webber, B., (Eds.), New York, 1981.

3. Cohen, P. R., & Perrault, C. R. Elements of a plan-based theory of speech acts. Cognitive Science 3, 1979, 177-212.

4. Cohen, P. R. On knowing what to say: Planning speech acts. Ph.D. Th., University of Toronto, Toronto, January 1978. Technical Report No. 118, Dept. of Computer Science.

5. Donnellan, K. Reference and definite description. The Philosophical Review 75, 1960, 281-304. Reprinted in Steinberg & Jacobovits (Eds.), Semantics, Cambridge University Press, 1971.

6. Fikes, R., & Nilsson, N. J. STRIPS: A new approach to the application of theorem proving to problem solving. Artificial Intelligence 2, 1971, 189-208.

7. Grice, H. P. Meaning. Philosophical Review 66, 1957, 377-388.

8. Grosz, B. J. Discourse Analysis. In Understanding Spoken Language, D. Walker, (Ed.), Chap. IX, New York City, 1978, 235-268.

9. Hintikka, J. Semantics for propositional attitudes. In Philosophical logic, J. W. Davis et al., (Eds.), Dordrecht-Holland, 1969. Reprinted in L. Linsky (Eds.), Reference and modality. New York: Oxford University Press, 1971.

10. Lewis, D. K. Convention: A philosophical study. Harvard University Press, Cambridge, Massachusetts, 1969.

11. Morgan, J. L. Toward a rational model of discourse comprehension. In Proceedings of Theoretical Issues in Natural Language Processing, D. Waltz, (Ed.), 1978.

12. Ortony, A., & Anderson, J. Definite descriptions and semantic memory. Cognitive Science 1, 1, 1977, 74-83.

13. Perrault, C. R., & Allen, J. F. A plan-based analysis of indirect speech acts. American Journal of Computational Linguistics 6, 3, 1980, 167-182.

14. Russell, B. On denoting. Mind 14, 1905, 479-493. Reprinted in D. Davidson and G. Harman (eds.) The Logic of Grammar. Encino CA: Dickenson, 1975.

15. Schiffer, S. Meaning. Oxford University Press, London, 1972.

16. Searle, J. R. Speech acts: An essay in the philosophy of language. Cambridge University Press, Cambridge, 1969.

17. Searle, J. R. Indirect speech acts. In Syntax and semantics, Speech acts, P. Cole & J. L. Morgan, (Eds.), New York, 1975.

18. Strawson, F. Intention and convention in speech acts. The Philosophical Review 73, 1964, 439-460. Reprinted in Logico-linguistic papers, London: Methuen & Co., 1971.

19. Webber, B. L. A formal approach to discourse anaphora. BBN Report 3761, Bolt Beranek and Newman, May, 1978.

20. Whitehead, A. N. The Concept of Nature. Cambridge University Press, London, 1920.

Official Distribution List
Contract N00014-77-C-0378

	Copies
Defense Documentation Center Cameron Station Alexandria, VA 22314	12
Office of Naval Research Information Systems Program Code 437 Arlington, VA 22217	2
Office of Naval Research Code 200 Arlington, VA 22217	1
Office of Naval Research Code 455 Arlington, VA 22217	1
Office of Naval Research Code 458 Arlington, VA 22217	1
Office of Naval Research Branch Office, Boston 495 Summer Street Boston, MA 02210	1
Office of Naval Research Branch Office, Chicago 536 South Clark Street Chicago, IL 60605	1
Office of Naval Research Branch Office, Pasadena 1030 East Green Street Pasadena, CA 91106	1
Office of Naval Research New York Area Office 715 Broadway - 5th Floor New York, NY 10003	1
Dr. Martin Epstein National Library of Medicine Bldg. 38A, 8th Floor Lab 8600 Rockville Pike Bethesda, MD 20209	1

Director, National Security Agency 1
Attn: R53, Mr. Glick
Fort G.G. Meade, MD 20755

Naval Research Laboratory 6
Technical Information Division
Code 2627
Washington, D.C. 20380

Naval Ocean Systems Center 1
Advanced Software Technology Division
Code 5200
San Diego, CA 92152

Dr. A.L. Slafkosky 1
Scientific Advisor
Commandant of the Marine Corps (Code RD-1)
Washington, D.C. 20380

Mr. E.H. Gleissner 1
Naval Ship Research & Development Center.
Computation & Mathematics Dept.
Bethesda, MD 20084

Captain Grace M. Hopper 1
NAICOM/MIS Planning Branch (OP-916D)
Office of Chief of Naval Operations
Washington, D.C. 20350

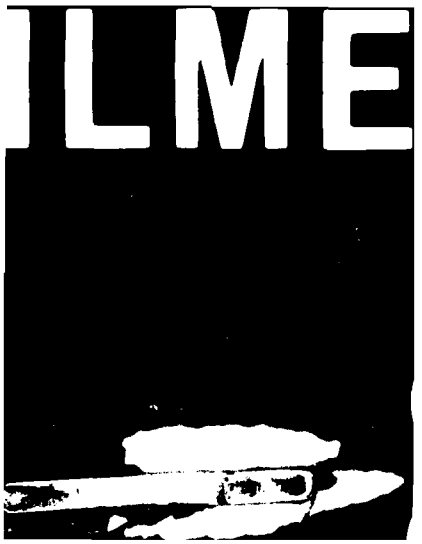
Mr. Kin B. Thompson 1
NAVDAC 33
Washington Navy Yard
Washington, D.C. 20374

Advanced Research Projects Agency 1
Information Processing Techniques
1400 Wilson Boulevard
Arlington, VA 22209

Captain Richard L. Martin, USN 1
3311 W. Avenue
Newport News, VA 23607

Director 1
National Security Agency
Attn: R54, Mr. Page
Fort G. G. Meade, MD 20755

Major James R. Kreer 1
Chief, Information Sciences
Department of the Air Force
Air Force Office Of Scientific Research
European Office of Aerospace Research
and Development
Box 14
FPO New York 09510


ILME