DOCUMENT RESUME

ED 136 780 IR 004 567

AUTHOR Levin, James A.; Archbold, Armar A.
TITLE Working Papers in Dialogue Modeling, Volume 1.
INSTITUTION University of Southern California, Marina del Rey. Information Sciences Inst.
SPONS AGENCY Advanced Research Projects Agency (DOD), Washington, D.C.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO ISI/RR-77-55
PUB DATE Jan 77
CONTRACT N00014-75-C-0710; NR-154-374
NOTE 82p.; For Volume 2, see IR 004 568

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS *Artificial Intelligence; Communication (Thought Transfer); *Computational Linguistics; Computer Oriented Programs; Computers; Dialogue; *Language Research; Man Machine Systems; *Models; Programing Languages
IDENTIFIERS Dialogue Modeling

ABSTRACT

The five technical working papers that comprise the two volumes of this document are related to the problem of creating a valid process model of human communication in dialogue. In Volume 1 both papers consider reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The second paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena, and for detecting their presence. (Author/WBC)

ARPA ORDER NO. 2930
NR 154-374

ISI/RR-77-55
January 1977

James A. Levin

Armar A. Archbold

Working Papers in Dialogue Modeling, Volume 1

2

INFORMATION SCIENCES INSTITUTE

UNIVERSITY OF SOUTHERN CALIFORNIA

4676 Admiralty Way/ Marina del Rey/ California 90291
(213) 822-1511

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ISI/RR-77-55	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) WORKING PAPERS IN DIALOGUE MODELING - Volume 1		5. TYPE OF REPORT & PERIOD COVERED Research 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James A. Levin, Armar A. Archbold		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0710
9. PERFORMING ORGANIZATION NAME AND ADDRESS USC/Information Sciences Institute 4676 Admiralty Way Marina del Rey, California 90291		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N RR042-06-01 RR042-06 NR154-374
11. CONTROLLING OFFICE NAME AND ADDRESS Cybernetics Technology Office Advanced Research Projects Agency 1400 Wilson Blvd., Arlington, VA 22209		12. REPORT DATE January 1977 13. NUMBER OF PAGES 80
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Personnel and Training Research Programs Office of Naval Research - Code 458 800 No. Quincy St. Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

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)

artificial intelligence, cognitive psychology, computer dialogue, linguistic comprehension, process model, natural language processing, reference.

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

(OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

The five technical working papers that compose this document (which appears in two volumes) were prepared as part of the Dialogue Modeling Project at ISI. Though diverse in scope, all are related to the problem of creating a valid process model of human communication in dialogue. All are unpublished and all but one are in a form intended for internal use by the project team; however, they are of interest beyond the boundaries of the project and have implications for related work in modeling human communication.

In Volume I both papers are on reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The other paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena and for detecting instances of them.

In Volume 2 the first paper concerns study methodology. It raises some of the following issues: how to choose between system-building and process-building, why studying cases is preferable to implementing general language-use functions as programs, how to control ad-hocness of results, why it is important to orient toward communication phenomena (in contrast to form phenomena) when studying natural language. The second is a design paper on the Match process of the Dialogue Model System, exploring methods for making it efficient and selective in its actions. The third concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation.

Each working paper appears with its original abstract or introduction.

ARPA ORDER NO. 2930
NR 154-374

ISI/RR-77-55
January 1977

James A. Levin

Armar A. Archbold

Working Papers in Dialogue Modeling, Volume 1

INFORMATION SCIENCES INSTITUTE

UNIVERSITY OF SOUTHERN CALIFORNIA

4676 Admiralty Way/Marina del Rey/California 90291
(213) 822-1511

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.

TABLE OF CONTENTS

ABSTRACT

The five technical working papers that compose this document (which appears in two volumes) were prepared as part of the Dialogue Modeling Project at ISI. Though diverse in scope, all are related to the problem of creating a valid process model of human communication in dialogue. All are unpublished and all but one are in a form intended for internal use by the project team; however, they are of interest beyond the boundaries of the project and have implications for related work in modeling human communication.

In Volume 1 both papers are on reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The other paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena and for detecting instances of them.

In Volume 2 the first paper concerns study methodology. It raises some of the following issues: how to choose between system-building and process-building, why studying cases is preferable to implementing general language-use functions as programs, how to control ad-hocness of results, why it is important to orient toward communication phenomena (in contrast to form phenomena) when studying natural language. The second is a design paper on the Match process of the Dialogue Model System, exploring methods for making it efficient and selective in its actions. The third concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation.

Each working paper appears with its original abstract or introduction.

Working Papers in Dialogue Modeling

Volume 1 - Section 1

Process Models of Reference

James A. Levin

Original Date: November 20, 1975

Table of Contents

One of the challenges in modeling human language capability is the fact that words are used to convey different meanings in different contexts. For example, the meaning of a pronoun is heavily dependent on the context in which it occurs, since it is used to refer to some entity mentioned elsewhere and has the same referent as this other expression. The meaning of instances of "it" may change radically from occurence to occurence, even within the same utterance. In the following portion of dialogue * :

"O: Did Runoff produce any output?
L: I dont know but I thought *it* would ask me what name *it* would be."

In the second utterance, the two "it"s refer to completely different things expressed in the first utterance, the first to "Runoff" and the second to "output".

Many of the process models of language understanding have dealt with this issue of repeated reference, in which one expression (often containing a pronoun or a determiner) derives meaning from some other expression in the discourse. Generally, these models contain a set of heuristic rules for ordering possible co-referential expressions and for testing them for appropriateness. For these models, the rules form a "bag of tricks" approach to repeated reference, rather than a systematic theory.

In this paper, we will examine in detail several models of language understanding that deal with repeated reference. After examining the operation of these models on a simple dialogue, we will classify the assembled "bag of repeated reference tricks" into two categories, and from this classification, propose a new approach for the process modeling of repeated reference.

1. Repeated Reference

Repeated Reference occurs when two sets of words in a discourse refer to the same concept. Let us call a set of words in an utterance defined as a unit on syntactic grounds an Expression. A Co-referential Expression is an Expression that has the same referent as some other Expression. Not all Expressions are Co-referential Expressions. Sometimes definite noun phrases are used to refer to generic concepts ("Do you know how *the mail system* works?"). Other times, an Expression can introduce a unique concept ("...*the ISI line printer*..."), or a concept that is unique in the given situation ("I just used a system called XOFF and it didn't give me *the normal output*"). Other Expressions in fact don't refer to anything ("*It's* 5 oclock.").

It is important to distinguish between Expressions and Concepts. Expressions are sets of words, while Concepts are the abstract entities which are the referents of Expressions (as well as referents of other non-verbal stimuli).

---

* Unless otherwise specified, all examples in this paper are taken from naturally occurring dialogues collected by the Dialogue Modeling Project at ISI. These dialogues are between a computer operator (labeled "O") and a computer user ("L"). The participants communicated remotely by typing into computer terminals using the TENEX "link" facility: whatever either person types appears simultaneously on both terminals.

We can easily detect Expressions because they are defined on a syntactic basis. In fact, we could use one of the existing parsing systems to mechanically detect Expressions. However, this isn't true for Co-referential Expressions. Some people (Baranofsky, 1970; Olney, 1969) have investigated the detection of Co-referential Expressions (sometimes called "anaphoric expressions") using certain syntactic aspects to distinguish the Co-referential Expressions from non-repeated ones. The language understanding systems described here all approach this detection problem by delaying the decision until it is trivial. They all look for preexisting referents for all Expressions, and those that have referents that are also the referents of other Expressions are then Co-referential Expressions.

2. Repeated Reference in Existing Systems

We will now concentrate on the heuristics for finding an existing referent, given an Expression. The operation of some of these will be illustrated on a simple dialogue, shown in figure 1. This is a real dialogue between a computer user and a computer operator typing over computer terminals using the "link" facility.

```
    LINK FROM [L], TTY 42

1.   1  2  3  4    5    6   7    8  9    10   11 12  13  14   15
L:  HOW OO I GET RUNOFF TO WORK, I KEEP XEQTN IT BUT IT JUST GRABS

1.  16  17    18  19   20   21   22  23   24   25 26   27     28
    MY INPUT FILE ANO THEN SAYS OONE BUT GIVES ME NO OUTPUT?  GA


2.   1    2      3    4  5   6   7     8
O:  THE OUTPUT COMES OUT ON THE LINE PRINTER


3.    1   2   3    4   5  6  7  8  9   10   11 12 13    14
L:  THROW IT AWAY BUT CAN I GET IT TO GE\EO TO A FILE?  GA


4.     1     2      3      4   5   6    7    8    9    10    11   12
O:  CONFIRM YOUR COMMANOS WITH A COMMA ANO YOU'LL BE QUERIEO FOR FILES,

4.  13    14
    ETC.  GA


5.   1    2
L:  THX MUCHO


    BREAK (LINKS)
```

Figure 1: Oialogue between a TENEX Operator ("O") and a User ("L") (The numbers and speaker labels have been added).

Our Dialogue Modeling Project at ISI has developed a procedure to gather annotations from observers of a dialogue transcript to use in evaluating dialogue models. Figure 2 shows the annotation produced by one observer of repeated reference phenomena in the dialogue in figure 1, using our repeated reference instructions (given in Mann, Moore, Levin & Carlisle, 1975). We will use these observations to evaluate the various repeated reference heuristics.

2.1 Repeated Reference in Verbworld

Verbworld (Rumelhart & Levin, 1975) is a recent language understanding system that deals with repeated reference in some detail.

There are two parts of this system where repeated reference is handled: the heuristics for handling pronoun reference (developed by Art Graesser) and the heuristics for handling definite determiners (developed by Donald Norman).

Pronouns

Third person pronouns (HE, SHE, IT, THEY, THEM, HIM, HER):
1. Look in the previous clauses for uses of this same pronoun.
If one is found, it is a Co-referential Expression.
else
2. Examine the noun phrases in the previous clause,
looking for a concept that matches the number and gender constraints.
Examine each clause in the following order: subject, object,
prepositional phrases. If a match is found, it is a Co-referential
Expression.
else
3. Start over with step 2.

Reflexive pronouns: (HIMSELF, HERSELF, ITSELF, THEMSELVES)
Examine previous noun phrases in this same clause for a concept with
the appropriate gender and number. If there is one, the phrase is a
Co-referential Expression.

Possessive pronouns: (HIS, HER, ITS, THEIR)
1. Examine noun phrases within the sentence, as in the case
of reflexive pronouns. If there is a match, it is a Co-referential
Expression.
else
2. Examine noun phrases in previous clauses, as in the case of third person pronouns.

Determiners

1. Evaluate the noun phrase. If the value is a unique concept, then
that value is the referent.
else
2. Evaluate any relative clauses in the noun phrase, and if this
results in an unique instance, this instance is the referent.
else

LINK FROM [L], TTY 42

L: How do I get RUNOFF to work, I keep executing it but it just grabs my input file and then says done but gives me no output? Go ahead

O: The output comes out on the line printer

L: Throw it away but can I get it to go to a file? Go ahead

O: Confirm your commands with a comma and you'll be queried for files, etc. Go ahead

L: Thanks much

BREAK (LINKS)

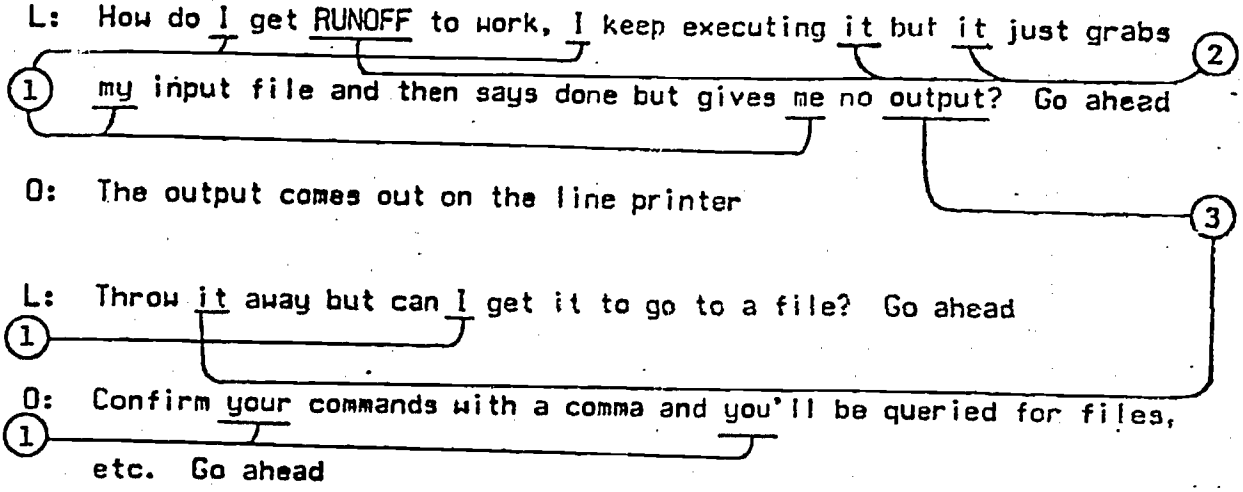

Figure 2: Dialogue annotated for Repeated Reference

3. Examine recent concepts mentioned in this and previous clauses and if there is a match with the features of the rest of the noun phrase, that concept is the referent. In particular, if an exemplar of the general concept was recently mentioned, that exemplar is the referent.

else

4. Assume that this determiner is being used in an indefinite sense, and create a new instance as the referent.

This system (like many of the others described below) was designed to be a participant in a dialogue rather than an observer of it. To enable it to operate as the observer of a dialogue, we must make one minor modification to the above heuristics: the first person pronouns ("I", "my", "me", "mine") take as their referent the present speaker; the second person pronouns ("you", "your", "yours") take the other participant as their referent.

Let us examine a hand-simulated operation of Verbworld on the dialogue in figure 1.

It would do well on all the personal pronouns in the dialogue, using the modified rule described above, assigning as referents the participants O and L.

In handling the "it" at 1.11, Verbworld finds "Runoff" (at 1.5) as its Co-referential Expression because (1) There are no "it"s in previous clauses, (2) The subject of the previous clause "I" isn't neuter and (3) The object "Runoff" is neuter.

For the "it" at 1.13, Verbworld finds the previous "it" at 1.11 and takes it as co-referential (and therefore , RUNOFF as its referent).

Both of these assignments correspond to the annotations by the observer (figure 2).

However, in turn 2, Verbworld runs into trouble with the definite noun phrase ("the output") at 2.1-2.2.. Depending on what referent the Expression "no output" 1.26-1.27 has, the heuristics will either find no repeated reference, or, worse, will find that expression ("no output") as co-referential because (1)"output" can't be evaluated to a unique concept, (2)there are no relative clauses, and (3)if "no output" is stored as a kind of output, it will fit the constraints of the Expression.

In turn three, the pronoun heuristics also run into trouble. The "it" at 3.2 is assigned "the output" (2.1-2.2) as co-referential. However, the annotations show "the output" as the generic class, and "it" (3.2) as an instance of that class.

Furthermore, the second "it" (3.8) is assigned the first "it" (3.2) as co-referential. The observer distinguishes these as different, as we can easily see from the semantics (You can't throw something away and then get that same thing to go to a file).

Finally, the observer annotated the generic "files", at 4.12 as a superset of the indefinite "a file" at 3.13. The Verbworld system would also make this assignment, because of the way the indefinite article "a" is defined.

2.2 Repeated Reference in SHRDLU

One of the most impressive language understanding systems to date is SHRDLU, developed by Winograd (1972). This system keeps track of "overall discourse", and looks in this overall discourse context for referent concepts. Winograd implemented a set of heuristics for repeated reference, most of which are summarized below:

General Tricks for Determiners

1. Find the set of all known objects that match the rest of the noun phrase.

then

2. If the right number are known, the object or the set of objects is the referent.

else

3. If there are too few objects, try to reparse the sentence (if this fails, print out a stored phrase asking for what the person meant).

or

4. If there are too many, try to find which were mentioned most recently. (if the right set of objects cannot be found, try reparsing as above, but with different error message about which were meant.)

General Tricks for Pronouns ("IT")

1. If there is another "IT" previously in the same sentence, it is a Co-referential Expression.

else

2. If there is another "IT" in the previous sentence, it is a Co-referential Expression.

else

3. For complex embedded NP's, check whether the "IT" is a reference to the NP it is in.

else

4. Look through previous clauses, looking for possible co-referential expressions, assigning plausibilities on the basis of the following:

a. Subject > Object > Prep Phrases
b. Main clause > Subordinate clause
c. "Focused" objects > non-focused

Choose the most plausible matching concept as the referent.

Special Case Tricks

IT: if used as propositional reference, the most recent action mentioned by the other participant is the referent.
THAT: if used as propositional reference, the most recent action mentioned by either participant is the referent.

I: tho roforent is :FRIEND
YOU: the referent is :SHRDLU

Now, despite the differences in SHRDLU's heuristics from those of Verbworld, it would produce the same referent resolutions on the figure 1 dialogue, incorrectly assigning the same referent to the "it"s in turn 3 and running into problems assigning "the output" in turn 2.

Why do these systems do so poorly on this simple dialogue? The Verbworld heuristics, for example, correctly found co-referential expressions for pronouns in 90% of randomly chosen text from an encyclopedia (Rumelhart & Levin, 1975). This disparity in performance gives us a hint of what the problem is. The encyclopedia contains only well-formed grammatical sentences, while the dialogue in figure 1 contains many ill-formed utterances, as is typical of real dialogue. Since most of the heuristics discussed so far are based mostly on surface syntactic features, it's not surprising that they do poorly on the syntactically ill-formed utterances in real dialogue.

Now, this shouldn't be taken as a claim that syntactically based heuristics are useless - only that there must be additional heuristics using semantic and pragmatic features that contribute to determining that Expressions are Co-referential.

2.3 Repeated Reference in Semantically Oriented Systems

Some systems find the referents for Expressions solely on the basis of semantic features, completely ignoring syntactic aspects. Quillian (1969) described such a system, and Reiger (1974) used a similar approach to do the repeated referencing for definite noun phrases within the MARGIE System (Schank, Goldman, Rieger & Riesbeck, 1973)

In this approach, a set of specifications for the referent are collected and a match process is conducted over the set of concepts in the current context (concepts that have been recently mentioned or used recently in some inference). A concept in context that matches the specifications is then a prime candidate for being the referent of the Expression.

Let us look at how a semantically based repeated reference system would deal with the dialogue in figure 1. As with the syntactic system, it handles the personal pronouns, given the modification to make the present speaker a specification of the referent of the pronoun "I", and the other participant a specification of the referent of "you".

For the "it" at 1.11, the partial specification of this referent is a non-human thing that can be XEQTed. ("XEQT" is an abbreviation for "execute", which is the action of starting a computer program). Given that XEQT was known to be an action applicable to programs, and that RUNOFF is known to be a program, this semantic approach could find RUNOFF as the referent.

There is a slight problem with the "it" at 1.13. For the correct referent to be found, the system would have to regard RUNOFF as an animate agent, since the referent is specified to be something which "grabs".

For the first Expression in turn 2, the semantic approach has the same problem as discussed for the syntactic approach here: it may find "no output" in turn 1 as co-referential.

The second Expression "the line printer" isn't a repeated reference, but the semantic approach may handle this correctly, depending on the exact comprehension and reference processes it has. If the knowledge by both participants about RUNOFF and TENEX systems includes line printers, and if this knowledge is accessed in understanding the dialogue to this point, then the correct referent for this Expression will be correctly found, even though it hadn't been mentioned previously.

In turn 3, the semantic system may see "the line printer" as co-referential to "it" on the basis of pure recency. The semantic system may be able to reject this, depending on what it knows about line printers and files. But this is a case in which syntactic clues would help a pure semantic system, since these clues would suggest "the output" as a better possibility.

There are cases in which ignoring syntax will cause a pure semantic system to fail.

"... there is a system that you can use that will let you manipulate your unsent mail. *It* is called mailstat, I believe. ..."

In this case, there are least two possible co-referential expressions to "it", "a system" and "your unsent mail", both of which fit the constraints. On syntactic grounds, we can see that "a system" is the co-referential expression. Clearly, we want to combine all the available syntactic, semantic, and pragmatic cues to repeated reference into one integrated approach.

3. Classification of the Bag of Repeated Reference Tricks

Let us look at a set of repeated reference tricks, shown in figure 3. These include the ones we have discussed already, plus additional repeated reference heuristics described in papers by Warnock (1972) and Baranofsky (1970).

We can classify the bag of repeated reference tricks into two categories, aspects of the possible Co-referential Expressions and aspects of the current Expression.

1. Proximity: The closer an utterance containing another Expression is to the current Expression, the more likely the other is to be Co-referential with the current one.
"I can't find any documentation on the program. I have a tape here at Purdue and I can't figure out what format *it*'s in."
The "it" refers to the tape, not to the documentation or to the program.

2. Syntactic role in the sentence:
2a. Subject/Object/Preposition phrases: The syntactic subject of an utterance is more likely to be an Co-referential Expression than the syntactic object, which is more likely than the preposition phrases.
"O: The output comes out on the line printer.
L: Throw *it* away ... "
The "it" refers to the output (the syntactic subject of the first utterance) rather than to the line printer (the syntactic object).

2b: Superordinate/subordinate: Concepts expressed in a superordinate clause are more likely to be Co-referential that those in any subordinate clauses.
" ... the tape that file is archived on seems to be a bad tape. We can't seem to get *it* to read ... "
The "it" refers to the tape rather than to the file (in the subordinate relative clause).

2c: Topicalization: Some special syntactic constructions (such as cleft sentences) can be used to emphasize one element of an utterance. These emphasized Expressions are more likely to be co-referential than unemphasized ones.
" ... there is a background job running here that checks to see if there is any unsent mail. Once *it* finds some, *it* tries to resend it."
The two "it"s are co-referential with the topicalized "background job" rather than with the "unsent mail".

3. Centrality: an concept which has previously been referenced more than once is more likely to be referenced again than one referenced only once.
"L: ... Any chance I can recover [file name] from the most recent system dump?
O: Probably, let me look for *it* and get back to you, ok?
L: Could you SNDMSG to me, one way or the other? I won't be doing anything about *it* tonight. If *it* is there, I will be forever grateful to recover *it*."
The "it"s by L all refer back to L's file, rather than to the system dump or the one way or the other, at least partially due to the previous reference by O.

4. Current topic: An Expression which refers to a concept in the current topic is more likely to be Co-referential. Deutsch (1974) observed that repeated reference can normally be made only to concepts that are part of a currently open topic. Once a topic is closed, it must be reopened before concepts within it can be referenced again.

"L: ... Can you recover those files for me..as far as I know they were in the directory on the 16th...the names are ...

[intervening dialogue]

O: OK I have found *the files you want* I will retrieve *those* for you ... "

In the second utterance, O had to initially specify the files in some detail, but once the topic was re-established, she could use just "those".

Figure 3a: Aspects of Possible Co-referential Expressions

1. Specification by the current Expression:
   1a: The the pronoun or determiner in an Expression often specifies that the referent concept be a particular number and/or gender.
   "How do *I* get Runoff to-work?"
   The pronoun "I" completely specifies the referent to be the speaker.

   1b: The other words in the current Expression often further specify the referent.
   "I have found *the files you are concerned about* ... "

2. Specification by the verb in the same utterance: The verb in the same utterance as the Expression often specifies that the referent have certain properties.
   "Throw *it* away."
   The verb (and verb particle) "throw away" specifies the referent to be something of little value to the speaker.

3. Specification by the whole clause containing the current expression: Sometimes the particular combination of a verb and its noun phrase arguments puts constraints on the referent.
   "Did *it* produce any output file?"
   The combination of the verb and the object limit the referent to being a computer program of some kind.

Figure 3b: Aspects of the Current Expression

3.1 The Possible Co-referential Expressions

The heuristics that deal with aspects of the possible co-referential expressions are a diverse lot, covering both syntax and semantics. One thing seems to be a common feature - each seems to reflect contributions to the salience of the various concepts in awareness. Some of the heuristics capture what is called "focus"; others reflect the fact that concepts in awareness are temporary, disappearing if they aren't repeatedly refreshed.

3.2 Aspects of the Current Expression

Initially, the referent of an Expression is completely unspecified . Each of the repeated reference heuristics given in the figure 3b can be seen as contributing specifications to this unspecified concept. Each heuristic may operate independently in adding its constraints to the referent of the current expression. And the end result is the partially specified referent.

This classification of the repeated reference heuristics into these two categories is straightforward. However, it suggests a general approach toward modeling repeated reference abilities in process models. There are two parts of a model - all those processes that contribute specifications to the referent of the current expression, and the processes that affect the salience of all the other currently active concepts.

4. An Activation Model of Repeated Reference

Let us assume we have some standard parser, that takes an utterance and chunks it up into Expressions.

Let's put each of these units into one place, called a Workspace, and give it a numeric value, called its Activation Rate. This Activation Rate reflects the momentary salience of the unit (it is similar to the "importance" metric discussed by Warnock (1972)). When a comprehension process attempts to put an Activation in the Workspace (to represent some new intermediate result), the Activation Model will first look for an identical existing activation. If none exists, the Model will create a new Activation with a specified Activation Rate. However, if an existing Activation is found, the Model instead increments its Activation Rate by the specified amount.

There is a Threshold value for existence of an Activation. Those Activations with rates below the Threshold are removed from the Workspace.

One part of the Activation Model is a set of rules for modifying the Activation Rates, derived from the heuristics in figure 3a. These rules are given below. The contents of the Workspace will be undergoing continual change, with new activations being created, and existing activations being deleted whenever they fall below Threshold. The contents at any one moment serves as the current context for reference resolution.

Let us explore the operation of this Activation Model of Repeated Reference by hand-simulating its performance on the dialogue in figure 1. To do this, we first have to specify a number of parameters (in a somewhat ad hoc manner) for the rules given

below. The first part of each rule description covers general aspects, and the second part gives a precise specification (with ad hoc parameters) that will be used to show the operation on the example discourse.

4.1 Rules for Modifying the Activation Rates of Possible referents

1. Recency of utterance rule:
   As each new utterance comes in, decrement the rates of all the preexisting activations.
   For the example below, decrement all existing activations by 50%.

2. Primacy within utterance rule:
   Increment the activation rate of the concept representing each unit of an utterance as it enters, but by a successively smaller amount. A subclaim of this Model is that the three separate syntactic factors 2a (Subject/object/preposition phrase), 2b (Super/subordinate clauses), and 2c (Topicalized constructions) in figure 3a can be captured in this one rule.
   For the example, increment the activation for the first expression by 10 (arbitrary) units, the next by 9, etc.

3. Centrality:
   When a concept has been referenced more than once, the same activation will have been incremented by each reference. Thus it will be more salient than if referenced only once, and so more likely to be selected again as a referent.
   The centrality aspect is already captured by the Activation Model itself, so we need no separate rule.

4. Current topic:
   Given a comprehension process for detecting topic structure, a rule that incremented the activation rates of all components of a newly detected topic would capture the current topic aspect. We don't yet have a model of this process, so we can't use this rule. But this illustrates the way that progress in modeling other aspects of natural language can be easily interfaced to this Activation Model extending its capabilities.
   The operator-linker dialogues, such as the one in figure 1, generally contain only one topic throughout, so the present lack of a topic structuring process isn't very noticeable for reference resolution in the cases we have been considering from these dialogues.

Now what do we do with an Expression? Well, we treat it much like any other unit. We create an activation for it, and start constraining what can fit in this spot by adding specifications. In the particular repeated reference model we are developing here, there are a set of rules for adding to this specification of the Expression, which are derived from the heuristics described in figure 3b.

So far, we have specified the aspects of possible co-referential expressions (figure 3a) as a set of rules for modifying activation rates within the Activation Model. Now we have to specify the other part of the model, capturing the aspects of the current expression (figure 3b).

4.2 Rules for Specifying the Referent

1. Definitions of the reference words:
   The reference words (pronouns & determiners) themselves contribute specification of number and/or gender.
   For the example, we need to define only the following words, so that these specifications can be added to the referent of the expression that they are in.
   "I" or "me" or "my" --> singular and human and current speaker
   "you" or "your" --> human and current hearer
   "it" or "this" --> singular
   "these" or "those" --> plural

2. Explicit modifiers in Expressions:
   If there are other words in the current expression, these further specify the referent. These include adjectives and nouns, prepositional phrases, and subordinate clauses.
   The following are words which are in expressions in the example dialogues and which add specifications to the referent: input, file, output, line printer, commands, comma.

3. The verbs in the same clause as the current expression:
   The kind of action described by the verb often adds further specifications to the referent.
   These are the verbs that occur in the example dialogue, most of which add specifications to their arguments: get, keep, xeqt, grab, give, come out, throw away, confirm, query.

4. The clause that the current expression is in:
   The event described by the rest of the clause also helps specify the referent.

The model then applies a match process between the partial specification of the referent of the current expression and the salient concepts in the Workspace, and selects the most salient concept that best matches the specification.

Let us now examine a hand-simulated operation of this Activation Model on the dialogue in figure 1.

For all the instance of first person pronouns ("I" at 1.3, 1.8, 3.6; "my" at 1.16; "me" at 1.25), the referent will be specified to be the current speaker and the expression will acquire that person as the referent.

The second pronouns ("you" at 4.8; "your" at 4.2) will all acquire the current hearer at turn 4 (the person L) as their referent and thus be co-referential expressions.

Let us focus on the "it"s, since these are the most challenging cases. When the "it" at 1.11 occurs, there are only a few activations in the workspace: person L (rate

4.5) and Runoff (rate 4). "It" specifies the referent to be singular; the verb specifies the referent to be a program. The only match is with Runoff, so that is acquired as the referent concept.

For the "it" at 1.13, the workspace now contains the same activations, but with different rates: person L (rate 7.25) and Runoff (rate 6.5). The specification of the referent are singular, animate, and involved with input files. Again, Runoff is acquired as a referent, making the expressions at 1.3, 1.11 and 1.13 all co-referential. This agrees with the annotations of the observer (figure 2).

Things are a little more complicated for the "it"s in turn 3, since there are more activations in the workspace at that time. These are the activations: specific output (rate 5), specific line printer (rate 4.5), person L (rate 3.3), non-existent output (rate 2), the word "done" (rate 1.1), Runoff (rate 1.9), and specific input file (rate 0.8). The specification of the referent include singular, movable, valueless to person L. The best matches are with the specific input and the specific output, and since the specific output is more salient, it would be acquired as a referent. This differs slightly from the observer's annotation, since he noted "it" as referring to a specific instance of the concept referred to by "the output".

Finally, the referent of "it" at 3.8 will be specified as being singular, movable, and an entity internal to computers. The state of the Workspace will be similar to that for the previous "it", with the Activation rates of the specific output higher and all others lower. The "it" would acquire the specific output as a referent. At one level the two "it"s in turn 3 are co-referential, but their low level referents are different and at this level they are not co-referential. We will examine this issue in more detail in section 7.

5. Complex Repeated Reference

So far, we have been primarily concerned with simple repeated reference, in which the co-referential expressions are simple noun phrases with relatively concrete concepts as referents. However, Expressions are often used to refer in much more complex ways. Our Dialogue Modeling Project at ISI has investigated two kinds of complex repeated reference. Text Reference and Propositional Reference (Mann, et. al., 1975; Archbold, 1975).

People sometimes use Expressions to refer to words or phrases that have been previously said. For example, we can talk about the last sentence or about this sentence - two text references. People more commonly use Expressions to refer to whole propositions that have been uttered previously. For example, I can refer to the previous description of repeated reference tricks - a reference to a concept that spans several pages.

These two kinds of references and their attendant complexities are discussed at length by Archbold (1975). To illustrate some of the many levels of text and propositional reference, consider the following (constructed) examples:

1. Four. That rhymes with score.
2. Four. That is a four letter word.
3. Fourscore and seven years ago. That's eighty seven years ago.
4. Four score and seven years ago. That's the opening phrase of a

famous speech given in 1863.
5. Four score and seven years ago. That was 1776.
6. Four score and seven years ago. That was 1888.

(See Archbold (1975) for a set of real examples that make this same point.)

Basically, people are able to reference a whole spectrum of concepts, including those concepts that represent intermediate results in the comprehension process. Among these are: a particular feature of the pronunciation of words, a specific use of a word, the word in general, a phrase, clause or sentence. These are all clumped under the category of text reference.

At a deeper semantic level, people are able to reference the concepts that represent various levels of comprehension and understanding. These include: actions, simple low level propositions, and larger scale units like topics. These are all types of proposition reference, and, like text reference, the referents are intermediate results of a comprehension process.

5.1 Complex Repeated Reference and the Activation Model

Existing models of repeated reference have had difficulty with these more complex kinds of reference. The few systems that address propositional reference at all handle it in an extremely limited and ad hoc way (for example, Winograd's SHRDLU), and none tackle the problem of text reference.

What about the Activation Model presented here? For this model to find a referent, the concept has to be in the Workspace. If we can get our language comprehension processes to put all their intermediate results into this Workspace, then these results will be temporarily available as possible referents of a text or propositional Expression. The Workspace will contain the whole spectrum of currently active concepts, all of which will fade away if not referenced again soon.

This way of expanding the scope of the activation model fits very naturally within a recently proposed general framework for process models (Levin, 1975). In this framework, called Proteus, all processing, at all levels, takes place within such a Workspace. However, even with more conventional language comprehension models, copies of intermediate results can be added to the Workspace, thus broading the scope of the referential processes.

5.2 Quotes

Quotes generally play a big role in much of the philosophical discussion of text reference. They are generally interpreted as signifying that the word itself is meant, rather than the underlying concept for which the word is the name.

However, in real dialogue, quotes (which occur rarely) are used in a broad variety of ways. (See the discussion of many of these by Archbold (1975).) One reaction to the hodge-podge of actual usage of quotes is to retreat to the performance/competence distinctions of linguists. However, once we take the actual uses seriously, we can see

that quotes are used as a "warning signal" to the hearer that the interpretation of the quoted word or phrase is meant to be different from the usual interpretation. Now, as we have seen in our examination of text and propositional reference, there are concepts at many different levels in the comprehension of a word or phrase that can be referred to. So, as a first approximation, we can extend the Activation Model of Reference to deal with quotes by suppressing the initial interpretation of a quoted word or phrase, thus allowing some less salient concept to be selected as the referent. This hypothesis for modeling quotes thus approximates the wide variety of ways in which quotes are actually used.

6. Non-Repeated Reference

People often use Expressions to introduce concepts not discussed previously. One way in which we do this is to refer to concepts not explicitly mentioned, but which are closely related to those that were. To account for this use of Expressions, Chafe (1972) introduced the notion of "foregrounding", in which the mention of a concept made closely related concepts available for referencing.

A classic (constructed) example of this is: "I rode a train today. I was allowed to toot the whistle". The phrase "the whistle" is a foreground reference. We don't just want to fall back on our default action of using the specification of the referent of the current expression (that this is some whistle that can be tooted). Instead, we are able to further determine that this is a specific whistle, which is controlled from the engine cab, etc. The first sentence "foregrounded" the knowledge about trains, so that the reference to "whistle" in the second can be determined to mean a very particular kind of whistle.

There are two systems which allow kinds of foreground references, Rieger's inference component of the MARGIE system (Rieger, 1974), and the SAM system (Schank, 1975).

In the MARGIE system, Expressions were detected by the parser and passed to the inferencing system for resolution. The set of possible referents considered by the inference system included not only those directly derived from previous utterances, but also those derived from any inference made from these utterances. In this way, the "inferred" concepts were "foregrounded" and thus available for referencing.

The SAM System (Schank, 1975) is an implementation using Scripts (Schank & Abelson, 1975) as a high level organization for language understanding .

In comprehending a particular set of utterances, a Script is found and used to guide further comprehension. Since a Script is an organized body of knowledge, the system can use it to generate expectations of future utterances. More interesting for this discussion, it provides a set of "closely related" concepts, all available for foreground reference.

6.1 Foreground Reference and the Activation Model

The Activation Model of Reference can be extended to deal with foreground reference in the same way as we extended it to deal with complex repeated reference. We put all the intermediate results of our inferencing and comprehension processes into the Workspace. For example, if we use higher level organizing concepts in comprehension, then whenever these concepts are used in comprehending an input, they will be put into the Workspace, and thus all the components will be available for referencing.

This proposed solution illustrates (and utilizes) the dependence of reference processes on other comprehension processes. Whenever some new comprehension process is developed, the capabilities of the referencing processes will also be expanded if the "results" of the new processes are added to the Workspace.

6.2 Failure to Find a Referent

What if we still don't find a referent concept? We already have a partial specification of the referent, and in many cases, this is all we need to know about this concept. Many Expressions seem to need no definite referent at all for comprehension to proceed satisfactorily. These cases are explored by Martin (1975) in some detail. In considering these cases, he developed a generalization of the notion of a simple pointer to a concept as the "referent" of an expression. These are "descriptions" with varying degrees of detail. These descriptions, which are partially specified referents, can be utilized in performing inferences, and also stored as knowledge known about the concept.

7. Referents and Reference

So far in this paper, the term "referent" has been used 128 times. Let us now examine explicitly how this term is being used, and therefore, what position on reference has been implicitly assumed.

We have talked about Expressions as sets of words which have concepts as their referents. One might be tempted to claim that the referent concept for an expression is the "meaning" of that expression. However, we have seen several ways in which this has to be modified. First of all, the particular concept which is the referent of an expression depends on the current context - the same expression can have two different referents in two different contexts. Given this observation, one might be tempted to view the referent of an expression as entirely context-dependent, and therefore that expressions by themselves have "no meaning".

However, our further investigation of complex reference and non-repeated reference have led us to a possible reconciliation of these two views of reference. Instead of a single referent concept, we found that an expression has a whole family of referents at many different levels. As the comprehension of a given expression proceeds, intermediate interpretations of the expression are generated, at successively more abstract levels. The initial referents represent the surface characteristics of the expression - the set of component letters, the shape or sound of the words. Successive levels include the low level semantic specifications - number, gender, the superset concepts.

The lower level referents of a given expression will be the same across context, while the more abstract referents will differ from context to context.

Now, we can consider two expressions to be co-referential whenever they have identical referent concepts at some level. That is, if expressions E1 in context C1 has referents (R11, R12, R13, ... , R1N) and expression E2 in context C2 has referents (R21, R22, R13, ... , R1N), then the two expressions are co-referential because they have the same referent concepts at level three and beyond.

For example, "Marina del Rey" and "*Marina del Ray*" are co-referential a a fairly low level; "La Jolla", "the place where UCSD is located", and "The Jewel of the Pacific" are co-referential at a slightly higher level, and "it" and "Runoff" are co-referential at a higher level in the utterance "How can I get *Runoff* to work? I keep xeqtn *it* ...".

Co-referential expressions are thus "the same" above a certain level, but different below that level. This difference is the reason why statements like "La Jolla is The Jewel of the Pacific" aren't empty tautologies (X is X).

A referent concept is then one of the family of concepts that represent an expression at some level.

8. Summary

In this paper, we have examined a number of the repeated reference heuristics used by language understanding systems. After observing the difficulties that existing models have with repeated reference in real dialogues, we divided these heuristics into two categories, those dealing with aspects of the possible co-referential expressions, and those dealing with aspects of the current expressions. Working from this categorization, we were able to propose an Activation Model for simple repeated reference, and then to extend it to text and propositional repeated reference, and finally to non-repeated reference. Some general issues of reference were examined in light of the Activation Model.

Acknowledgments

I want to thank the other members of our Dialogue Modeling Project at ISI, particularly Jim Moore, who read previous drafts of this paper and helped me clarify the terminology and concepts used here. Some of the ideas in this paper originated in discussion with our summer members, especially Alex Archbold (from Harvard) and Paul Martin (from Stanford).

References

Archbold, A. A., "Text Reference And Repeated Propositional Reference: Concepts And Detection Procedures", ***Working Papers in Dialogue Modeling - Vol. 1***, USC/Information Sciences Institute, ISI/RR-77-55, January 1977, Sec. 2.

Baranofsky, S. "Some heuristics for automatic detection and resolution of anaphora in discourse", Unpublished M. A. thesis, U. of Texas, 1970

Chafe, W. L. "Discourse structure and human knowledge". In R. O. Freedle & J. B. Carroll (eds.) ***Language comprehension and the acquisition of knowledge***, Washington, D.C.: Winston, 1972.

Collins, A. M., & Quillian, M. R. "How to make a language user". In E. Tulving & W. Donaldson (eds.) ***Organization of Memory*** New York: Academic Press, 1972.

Deutsch, B. G. "The structure of task oriented dialogs." ***Proceedings of the IEEE Symposium on Speech Recognition.*** Pittsburgh, PA., 1974.

Levin, J. A. "PROTEUS: An activation framework for cognitive process models." Doctoral dissertation, Univ. of Calif., San Diego, 1975.

Mann, W. C., J. A. Moore, J. A. Levin, and J. H. Carlisle, ***Observation Methods For Human Dialogue***, USC/Information Sciences Institute, ISI/RR-75-33, June 1975.

Martin, P. A. "Why describing is better than pointing." Unpublished paper. Information Sciences Institute, Marina del Rey, CA., October 1975.

Olney, J. C. "An investigation of English discourse structure with particular attention to anaphoric relationships". Santa Monica: System Development Corporation, TM-4229, 1969.

Quillian, M. R. "The Teachable Language Comprehender: A simulation program and theory of language". ***Communication of the ACM***, 1969, *12*, 459-476.

Rieger, C. "Conceptual memory: A theory and computer program for processing the meaning content of natural language utterances". Doctoral dissertation, Stanford Univ., AIM-233, 1974.

Rumelhart, D. E., & Levin, J. A. "A language comprehension system". In D. A. Norman, D. E. Rumelhart, & the LNR Research Group, ***Explorations in cognition***, San Francisco: W. H. Freeman, 1975.

Schank, R. C. "Using knowledge to understand." ***Theoretical issues in natural language processing***, Proceedings of an ad hoc conference, Cambridge, Mass., 1975.

Schank, R. C., & Abelson, R. P., "Scripts, plans, and knowledge," Paper presented at the Fourth international joint conference on artificial intelligence, Tbilisi, USSR, August, 1975.

Schank, R. C., Goldman, N., Rieger, C. J., & Riesbeck, C. "MARGIE: Memory, analysis, response generation, and inference on English", *Proceedings of the third international joint conference on artificial intelligence*, Stanford, 1973.

Warnock, E. H. "Anaphoric reference: Finding the antecedent". Unpublished paper, Harvard University, 1972.

Winograd, T. "Understanding natural language". *Cognitive Psychology*, 1972, 3, 1-191.

Working Papers in Dialogue Modeling

Volume 1 - Section 2

Text Reference and Repeated Propositional Reference: Concepts and Detection Procedures

Armar A. Archbold

Original Date: August 15, 1975

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
INTRODUCTION

I. INTRODUCTION

When people engage in dialogue, they quite frequently, at some point, refer back to and talk about something that has already been said in the dialogue, or refer forwards to something that is about to be said. In some cases, they may refer to a string of words uttered at some nearby point in the dialogue, thereby making a *"(repeated) text reference"*. In other cases, they may talk about some state of affairs, some statement, some belief, which is described, made, expressed or referred to nearby in the dialogue, thereby making a *"(repeated) propositional reference"*.

Both text reference and repeated propositional reference are of interest to the Dialogue Analysis Project. Two questions immediately arise concerning these dialogue phenomena, however. First of all, there is the conceptual problem: (a) is it possible to define these phenomena at all clearly? (b) if there are several different feasible definitions available, which should the Dialogue Analysis team choose, given its goals and interests? Secondly, there is the detection problem. Given a definition of these two phenomena, how can one determine what expressions may be involved in them, and which particular expressions are involved in a particular instance of them?

The present paper addresses these two problems. The first section will survey some philosophical and linguistic literature's treatment of the notions of "text" and "proposition" and of problems associated with these notions. The available notions will then be discussed and evaluated in the light of the team's interests and goals, and in view of some of the data that the team will have to account for. The second section will discuss various "clues" which might be relied upon to determine whether a given expression is being used in a dialogue to make a text reference or a repeated propositional reference. The focus will be upon the question: are there any operationalizable procedures for detecting pairs of expressions which are involved in repeated text or propositional reference?

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

SECTION I:

CONCEPTS OF TEXT REFERENCE AND REPEATED PROPOSITIONAL REFERENCE

In this section, the concepts of text reference and of repeated propositional reference will be examined, first as they are presented in some philosophical and linguistic traditions, and then in view of their operationalizability for the team.

Below, we shall use the following terminology and abbreviations:

REXPR	a referring expression, i.e. an expression which is used by a speaker in a dialogue to refer to, mention, or pick out an "object" or a set of "objects" (where "object" is taken in its widest possible sense, to include physical objects, people, states of affairs, events, actions, processes, abstract constructs such as the numbers or the quality of beauty, intentional objects such as thoughts or beliefs, hypothetical or fictional objects such as the child X and Y would have conceived together had they not separated, or Pegasus, and so on).
TREF	text reference
PREF	propositional reference
RREF	repeated reference: the use of two separate linguistic expressions to refer to (designate, mention, pick out, etc.) the same "object". The first referring expression involved we call "the antecedent referring expression" (AREXPR); the later of the two expressions involved we call "the consequent referring expression" (CREXPR).
RPREF	repeated propositional reference.

I.1 TEXT REFERENCE

I.1.1 NOTIONS OF TEXT REFERENCE IN THE LITERATURE

In the literature, the notion of text reference is closely bound to the distinction between the use of language and the mentioning of language. We ordinarily *use* language to talk about the world ; but we may also use language to talk about language, to discuss linguistic expressions - words, phrases, sentences, or whole texts or discourses. In the latter case, we ***mention*** or refer to linguistic expressions in order to say something about their phonetic, orthographic, syntactic, semantic, logical or pragmatic features. [1]

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION

THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

Traditionally it was thought that the use and mention of expressions were very distinct operations, and that to ignore the distinction was to run the risk of creating such nonsense as:

My dog contains three letters.

or

My "dog" is a hearty eater.

Though it was allowed that one could mention expressions by the use of proper names or definite descriptions, the traditional paradigm of text reference was text reference by quotation. A quotation-expression - i.e. a quoted expression plus its surrounding quote marks - is usually analysed as a name for the expression within the quote marks.

There is an important ambiguity in the notion of mentioning expressions. Text reference is not reference only to individual sounds or inscriptions. When we refer to linguistic expressions, we may refer either to tokens or to types. *Tokens* are particular, unique strings of marks or sounds. *Types* are not single homogeneous entities (there is no such thing as a type-word), but are rather classes of individual tokens, grouped together by reference to some set of taxonomic criteria [2], which can be referred to by the use of

[1] The fact that natural language can be used to talk about natural language expressions (to "talk about itself" in some sense) leads to well known antinomies. Consider the next sentence. The third sentence of footnote 1 of this paper is false. Is the preceding sentence true or false or both or neither? In order to avoid such paradoxes, logicians who have talked about linguistic expressions in order to define their truth conditions have distinguished between the language they examine - the object-language - and the language they use in their analyses - the meta-language. Cf. Alfred Tarski, "The Semantic Conception of Truth", in Leonard Linsky, ed., "Semantics and the Philosophy of Language", University of Illinois Press, Chicago, 1952.

[2] Two remarks are in order here. First, we may very well be unable to explicitly specify some or most of the criteria which we use to group tokens into classes, i.e. to decide when two different tokens are of the same type. Secondly, it is probably wrong to thing that we sort tokens into types by reference only to orthographic or phonic criteria, without any consideration of their semantic role. What we refer to when we perform a text reference to a token is a sound or a mark which belongs to a language, which was produced for a characteristic linguistic purpose, or was produced in a context in which describing the sound or mark in terms of its syntax, relative to a framework in which such events or marks can be systematically described, is an appropriate activity to engage in. And we disregard orthographic differences between tokens which would have no impact upon their semantic roles. Cf. D.V.C. Lincicome, "Systematically Ignored Differences and the Identity of Propositions", Foundations of Language, Vol 12, No. 1, September 1974, Section 2.

species-words similar to "dogs" (which refers to the class of all dogs). A convention which one might adopt to avoid confusion here is to refer to tokens by singling them out by means of ostension or definite description, and then christening them with a proper name. For example, the mark on the next line

animal

one could call Tom. Species-words can then be used to refer to classes of such individual tokens; species-words could be formed by spelling the token backwards, inserting dots between the letters.

Thus if the two marks on the next line

animal towards

are christened Dick and Harry, one could say that both Tom and Dick are l.a.m.i.n.as - that is, they are both tokens which are of the l.a.m.i.n.a class or species [3]. Given the distinction between tokens and types, one must clarify one's analysis of the role of quotation-marks, and separate out those cases in which quotation-marks form the name of a token, and when they form the name of a type to which the individual quoted expression token belongs.

Let us consider some examples of TREF. The examples below - all partly or wholly concocted for purposes of illustration - involve two or more underlined expressions which refer to the same text; such repeated reference is simply designed to highlight the TREFs being performed.

(1) A: The third letter from the top of the list on the blackboard is "D".
B: Yes, I know. It's written in red chalk. But I can't make out the letter directly below it.

(2) A: John Smith is calling himself "Hiroto Texagewish" these days.
B: His assumed name is certainly hard to pronounce!

(3) A: On checklist G/8-2 under step 5, there's a statement "disable all jets on two adjacent quads". Is that what you are talking about?
B: No, "disable all jets on two adjacent quads" is not what I was referring to.

(4) A: SHOULD I TYPE ATT LINKER (PASSWORD), JOB NUMBER?
B: YES, THAT'S WHAT YOU SHOULD TYPE IF YOU WANT TO RE-ATTACH.

(5) A: John gave me a good example of a tongue-twister yesterday. What was it now? Ah yes, I remember. How much wood would a wood-chuck chuck if a wood-chuck would chuck wood?.
B: Well, the tongue-twister he proposed was easier to pronounce than the one Mary came up with.

[3] L. Goddard and R. Routley, "Use, Mention and Quotation", The Australasian Journal of Philosophy, Vol. 44, No. 1, May 1966.

(6) A: You're coming in a little louder now.
B: Fred, you'll have to say that again. I couldn't hear you; there's too much background noise.
A: Okay. I said, you're coming in a little louder now.

Consideration of the above examples leads to the following remarks:

(1) The linguistic expressions which are used to refer to (mention) other linguistic expressions are of two types. The first type *exhibits* or *replicates* the expression to which it refers. Thus in the two following sentences

John wrote "Today's lesson is Chapter V" on the board.
Should I type TTY or TTY:?

the underlined expressions are used to refer to expressions which they replicate or exhibit. The difference between exhibiting and replicating can be shown by the following examples:

The following letter, "D", was not written by hand.
(exhibition)
Yesterday, John wrote "Today's lesson is Chapter V" on the blackboard.
(replication)

In the first case, one is exhibiting a token and talking about that very token, and no other. In the second case, one is talking about a token John produced yesterday, and one is doing so by exhibiting a token of the same type as the token he produced yesterday; in that sense one is replicating a token, by producing and exhibiting a "copy", a token of the same type.

The second type of linguistic expression which is used for TREF neither exhibits nor replicates the expression to which it refers. An example of such an expression would be "The sentence John wrote on the board yesterday".

(2) Exhibiting or replicating text referential expressions may or may not use quotation marks: compare the examples (1-3) and (4-6). This fact flies in the face of the traditional view according to which one only exhibits or replicates expressions to which one wants to refer by using quotation-marks, quotation-marks being seen as operators which form a proper name of the string enclosed within them. One can exhibit or replicate a linguistic expression not only by quotation, but also by capitalization, pauses, indentation or spacing, intonation and many other ways besides. There thus seems to be no one reliable orthographic sign of exhibition or replication; there may, however, be a list of such signs. In any case, the function traditionally assigned to quotation marks can be performed by many other marks or sounds. (One should also note that quotation marks are sometimes used in complex ways which bear only a faint resemblance to their stereotypical use, as when they are employed in ironical remarks, as shudder quotes or snigger quotes. Examples: "These days, one never gets price stability, only decreases in the rate of increase of inflation. That's "progress".", or Lenin's statement "We will "support" the Mencheviks as the rope supports the neck of the hanged man!")

(3) Text referential expressions which are not of the "exhibition" or "replication" types may be (a) pronouns or (b) noun phrases. The noun phrases may be proper nouns (think of referring to a string of marks on a piece of paper as "Tom"), (modified) common nouns ("a password", "the loud shout") with or without relative clauses.

(4) Text referential expressions which are not of the "exhibition" or "replication" type may be combined with those that are to form complicated text referring expressions. Consider example (4), where "(PASSWORD), job number" are used alongside two replication text expressions. It is interesting to note that the order in which these referring expressions occur is the same as the order in which the referents of each individual expression must occur in order to form a token of the type which the entire string "ATT LINKER (PASSWORD), job number" denotes.

One might conclude that text reference is a *pure* mention of phonetic or orthographic tokens or classes of tokens. However, the traditional clearcut distinction between use and mention has been attacked of late - and justifiably so. Things are not so simple.

> "There are ... many sentences in which an expression is both introduced and is also used; in particular, sentences which are used to convey both linguistic and factual information. This is especially so in sentences containing the words 'call', 'distinguish', 'determine', 'is called', ... 'satisfy', or compounds of such words. Consider, e.g. 'That sleek red-coated dog is Rover', 'The "Queen Elizabeth", which is so-named (so-called) after the present Queen Mother, sailed for Southampton yesterday', 'What is halva?', 'Call her a shrew',
>
> 'If triangles are taken as three-sided figures then they have...', and also indirect speech forms in which the speakers actual words are reported."[4]

There are sentences containing text references in which the text refered to must not only be considered as an uninterpreted orthographic object, but also be "read with understanding", i.e. interpreted. Some examples of such sentences are:

(1) The sign says, "George Washington slept here", but
I don't believe he ever did.

(2) Whenever Fred sighs "Boy, do I need a drink", he expects
you to fix him one.

(3) What he actually said was "It's clear that you've given
this problem a great deal of thought", but he meant
quite the opposite.

(4) "I talk better English than both of youse", shouted
Charles, thereby convincing me that he didn't.

In all of the above, the presence of pronominalization, ellipsis or semantic anaphora

[4] L. Goddard and R. Routley, op.cit., pg. 22.

involving terms both within and outside of the quotation marks shows that the quoted sentences must have been interpreted.[5] Thus if we say that text reference involves the mention of a linguistic expression, *we must allow that such mention may be accompanied by use, and that therefore we must count as text reference not only reference to tokens or types alone, but also reference to tokens or types paired with their meaning, some interpretation, the statement they were used to make on some occasion, or the proposition they express.*

I.1.2 THE TEAM'S INTEREST IN TEXT REFERENCE.

In light of the concepts introduced above, let us now consider the notion of text reference which the team is interested in.

In "Observation Methods for Human Dialogue", we find the following explanations of the concept:

> "...a so-called 'Text Reference', made to a string of words in the preceding dialogue itself (and not to the referent of that preceding string of words!)." (pg. 22)

> "A Text Reference occurs whenever reference is made to previously occurring words within the transcript. For example, in the sentence 'Go 3 blocks and turn North; by North I mean towards the mountains.' the second use [i.e. occurrence - A.A.] of "North" is a Text Reference to the first. We call this a Text Reference because it refers to the previous use [occurence] of the word itself, rather than to its meaning." (pgs 26-27)

These passages indicate that the team is interested only in singular text references to words previously used in the dialogue. Note that if we interpret these indications strictly, we must conclude that only TREF to previously used tokens are to be focused upon - since word-tokens, and not classes of word-tokens, are used at a particular time and place by participants in a dialogue. On this strict interpretation, the example given, viz.

Go 3 blocks and turn North;
by North I mean towards the mountains.

is of interest only if one reads the second sentence as equivalent to 'I intended to use the token of type "North" which occured in my previous sentence to mean towards the mountains.' It would not be of interest if the second sentence were read as equivalent to 'All tokens of type "North" which I utter I use to mean towards the mountains', for in that case the second occurence of "North" would refer not to the previous token but to a class of tokens of which the previous token is a member.

However, examination of the examples given in "Observation Methods in Human Dialogue" and discussion with team members indicate that they are concerned with a much wider variety of phenomena. What they are concerned with and what they wish to refer

[5] Barbara Hall Partee, "The Syntax and Semantics of Quotation", in S.R.Anderson and Paul Kiparsky, "A Festschrift for Morris Halle", Holt, Rinehart and Winston, New York, 1973.

to by the term *"Text Reference" consists of:*

(1) Singular text reference where an expression is used to refer to either (a) a token which is used either before or after it in the dialogue, or (b) a class of tokens of which one or more tokens used in the dialogue are members.

(2) Repeated text reference where the AREXPR and the CREXPR refer to either (a) the same token, or (b) the same type, or (c) a type and a token which is a member of that type, or, lastly, (d) two token-classes (i.e. types), one of which is a subset of the other. Either or both of the referents of the AREXPR and the CREXPR may either be or contain as a member a token used in the dialogue.

We may repeat here what we have stressed above: if an expression E1 is used in a dialogue, and is elsewhere referred to by means of another expression E2, E2 is a text-referring expression which is involved in a singular text reference (in the sense of (1)) if it is used to refer to E1 as a token, i.e. if it is used to refer to some inscriptional or orthographic features of E1. However, E2 may also (at the same time) be used to refer to the meaning, the referent, or some other non-textual feature of E1. Text reference does not exclude interpretation.

REPEATED TEXT AND PROPOSITIONAL REFERENCE: CONCEPTS AND DETECTION
THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

I.2.1 PROPOSITIONS AND PROPOSITIONAL REFERENCE

Whereas the notion of "text" is relatively unproblematical and unconfusing, given certain simple distinctions, the notion of "proposition" is a difficult one to formulate clearly. Unlike "text", "proposition" is very much a technical philosophical term. We find three main concepts of what a proposition is, formulated and used by (1) linguists, (2) logicians or philosophers primarily concerned with logical matters, and (3) by speech-act theorists.

1.2.1.1 Those linguists who make use of the term "proposition" often equate it with the *meaning, reading or semantic interpretation of a sentence*. Let us consider for example the Katzian tradition of transformational semantics. In this tradition, the meaning or semantic interpretation of a sentence is a set of sets of structured markers which are assigned to the sentence on the basis of (a) the semantic markers assigned to the component words by a dictionary, (b) the syntactic structure of the sentence, and (c) a set of semantic "projection rules". The semantic interpretation of a sentence is a theoretical construct which is adequate if in conjunction with the rules of a semantic theory can predict the semantic properties of sentences (such as synonymy, ambiguity, redundancy, presupposition, entailment, and so on). Each set of structured markings is a "reading" or a "proposition". If a sentence is assigned several sets of structured markings it is ambiguous and is said (according to the semantic theory which assigns the markings) to *express several different propositions*.

> "Sentences are frequently ambiguous, that is, they express more than one sense. Thus, we shall frequently say that an n-way ambiguous sentence expresses n distinct propositions. (We have taken the term 'meaning' to refer to the sum of the propositions expressed by a sentence ...). We also understand 'proposition' to convey what synonymous sentences have in common by virtue of which they are synonymous. Sentences that are synonymous on a sense [i.e. on one of their readings - A.A.] are thus said to express the same proposition, and fully synonymous sentences are said to express the same set of propositions. Semantically anomalous sentences express no proposition at all."[6]

Note that if one regards propositions as the 'readings' of sentences, one allows that not only declarative, but also interrogative, imperative and hortatory sentences express propositions.

1.2.1.2 The logicians' notion of proposition is distinct from the linguistic notion, at least prima facie. Logicians are concerned with formal constraints on inference, and with the notions of truth and falsity which are needed to account for formal validity of arguments. They have thus made use of a concept of "proposition" which is tantamount to the notion of a "*truth-vehicle*" :

> "On the logical account of propositions, propositions have been regarded, first and foremost, as truth-vehicles. That is, propositions are taken to be either the things or some of the things which are true or false. There are several

[6] Jerrold J. Katz, "Semantic Theory", Harper and Row, New York, 1972, pg. 120.

> motivations for this view of propositions. One historically important motive has been to provide a subject matter for logic, something for logic to be about. Logic is, in the first instance, the study of inferences. Whether an inference is valid or invalid depends neither on the particular subject of discourse nor on the determinate mode of linguistic expression but solely upon the formal relations between premises and conclusion. Propositions may thus be conceived as sorts of entities which stand necessarily in such relations as entailment and contradiction, and it is these relations which constitute the grounds of valid and invalid inference and are reflected in particular linguistic embodiments. ...
>
> A second motivation for the logical account of propositions is found in the classical correspondence theory of truth. On this view, truth is regarded as a relation between what is the case in the world, the facts, and the thing, whatever it is, which is true. Propositions have traditionally been cast in the role of the second term of this relation." [7]

Thus on the logicians' view, as on the linguists', there is a distinction between sentences - i.e. strings of inscriptions or sounds which belong to some language - and the propositions which sentences express. Thus the following distinct sentences (taken either as tokens or types)

(1) The moon is smaller than the sun.
(2) The sun is larger than the moon.
(3) La lune est plus petite que le soleil.
(4) Le soleil est plus grand que la lune.

would be said to express the same proposition, and to be true because they all express the same true proposition.

One may to some extent distinguish between a traditional logician's notion of a proposition, and more recent notions.

1.2.1.2.1 My own stereotypical characterisation of the traditional notion is as follows:

(1) a proposition is an abstract object;

(2) there are non-denumerably many propositions;

(3) propositions exist independently of language (ie. there are many propositions which are not, may never be, and perhaps could not be, expressed by a sentence;

(4) a proposition is something which is itself true or false in an absolute (timeless) sense;

(5) some propositions are expressed by declarative sentences;

[7] Rosenberg, Jay F., and Travis, eds., Charles, "Readings in the Philosophy of Language", Prentice-Hall, New Jersey, 1971, pgs.219-220.

(6) if two sentences express the same proposition, then they are analytically equivalent (ie. their meanings constrain their truth conditions in such a way that they are either both true or both false in any situation, state of affairs, or "possible world");

(7) however, if two sentences are analytically equivalent, they need not express the same proposition: analytical equivalence is less stringent a requirement than identity of proposition expressed (thus though 'John is a bachelor' and 'John is an unmarried adult male human being' are analytically equivalent, we do not want to have to say that they express the same proposition);

(8) if two sentences are logically equivalent, i.e. have the same truth-values in all models, then they need not express the same proposition (for example, 'x=x' and 'Fx or ~Fx' are logically equivalent, because true in all possible worlds, but we do not want to say that they express the same proposition);

(9) propositions are what are believed, doubted, hoped for, etc., i.e. they are the objects of belief, doubt, hope and the other so-called "propositional attitudes".

1.2.1.2.2 The traditional logical notion of proposition, though it can be made precise to some degree [8], has been much criticised as ontologically unnecessary and obscure [9]. Logicians are interested in what is true or false. What we usually term true or false are *sentences uttered by people in certain contexts and interpreted in certain ways*. It was argued that to postulate the existence of propositions above and beyond uttered sentences was simply to complicate furthur an already vexing question with such additional quandries as the exact nature of the relationship between propositions and the sentences which "express" them. It was felt that the real problem is to determine the nature of the dependency of the truth of sentences upon the context of their use:

> "A sentence is not an event of utterance, but a universal: a repeatable sound pattern, or repeatedly approximable norm. Truth cannot on the whole be viewed as a trait, even a passing trait, of a sentence merely; it is a passing trait of a sentence for a man. 'The door is open' is true for a man when a door is so situated that he would take it as the natural momentary reference of 'the door' and it is (whether he knows it or not) open. The individual event of utterance can still be described as true absolutely, since a time and a man are specific to it; but talk of sentences as true for men at times covers more ground, for it includes cases where the sentence is not uttered by the man in question at the time in question.

[8] Cf. Jan Berg, "What is a Proposition?", Logique et Analyse, Vol. 10, Dec. 1967 (summarized).

[9] Cf. the arguments summarized in Howard Pospesel, "The Non-Existence of Propositions", The Monist, Vol. 53, April 1969 (summarized).

> Relativity to times and persons can be awkward on account of the supplementary specifications in which it keeps involving us. This is no doubt one reason why philosophers have liked to posit supplementary abstract entities - *propositions* - as surrogate truth-vehicles."[10]

If a person utters a sentence in a certain manner in a certain place, at a certain time, to a certain interlocutor, in brief, in a certain context and "cotext" (verbal or dialogue context), he has said something, *made a statement* , by *using a sentence in a certain way* [11]. Many contemporary philosophers regard statements as "primary truth-bearers"; they regard statements as what are (timelessly) true or false. Some furthur claim that statements are representable by *eternal sentences* , i.e. sentences the values of all of whose indexical terms have been explicitly specified, whose truth-value consequently stays fixed through time and from context to context. One might thus think of a statement as a pair comprising a sentence and a complete interpretation of that sentence; the interpretation of the sentence would ideally provide a complete function from possible states of affairs (possible worlds) to truth-values for that statement. Such a function, in other words, would be a complete and precise specification of the truth-conditions of that sentence as used; it would include a set of specifications of truth-conditions which stem from the conventional semantic meaning of the sentence taken just as a sentence of the English (or other) language it belongs to, plus complete specifications of the values of all the indexical terms in the sentence, given the context of use, and unambiguous definite descriptions of the referents of the ambiguous or vague definite descriptions in the sentence. Such interpretations correspond to what some formal logicians have called intensions in their models. Such complete specification is possible by fiat in the domain of formal semantics. It is a moot point whether such a complete specification is possible for a natural language sentence uttered in everyday circumstances.

It is important to note that philosophers may speak of propositions being expressed by sentences, and of statements being made by uttering sentences, but they do not talk of sentences denoting or referring to propositions - except in one case. They do talk of nominalized sentences in modal or intentional contexts as being "proposition-denoting expressions". Consider the sentences "It is impossible that Mary is sick" and "John believes that Mary is sick". In both of these cases one finds the nominalized sentence "that Mary is sick". A traditional analysis of the logic of such sentences claims that the expression "that Mary is sick" is an expression which denotes the proposition expressed (but not denoted) by the sentence "Mary is sick" when occuring outside such "opaque" contexts as modal or (especially) intentional contexts.

[10] Willard V.O. Quine, "Word and Object", The M.I.T. Press, 1960, pgs. 191-192.

[11] For important articles which make use of this terminology, cf. E.J. Lemmon, "Sentences, Statements and Propositions", in J.F. Rosenberg and Charles Travis, eds., 'Readings in the Philosophy of Language', Prentice-Hall, Englewood Cliffs, New Jersey, 1971, and the two articles by P.F.Strawson - "On Referring", Mind, 1950, and "Identifying Reference and Truth-Values", Theoria, Vol. XXX, 1964. All of the three above-mentioned articles are summarized.

THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

The two differences between the linguistic and the logical notion of 'proposition' (or 'statement') which are most relevant to us are well summarized by Katz:

> "...[a] point about our [i.e. some linguists'] use of the term 'proposition'... is that the class of propositions cannot be identified with the class of statements, where statements are understood as the logical objects that are the bearers of truth values (i.e. as the objects that obey the law of the excluded middle). The fact that our characterization of the class of propositions encompasses a multitude of nonassertive propositions (questions, requests, etc.), for which it makes no sense to talk about truth and falsity, makes this amply clear. But we cannot even identify the class of statements with the class of assertive propositions, since a proposition with a token indexical element cannot have a fixed truth value."[Ibid., pgs. 122-123]

The third notion of 'proposition' which one can pick out of the literature is that of the speech act theorists. To put it rather vaguely, for speech-act theorists, the proposition expressed by someone who utters a sentence-token is *what is left, syntactically and semantically, in the uttered sentence token after all of its constituents relevant to a determination of its illocutionary force have been abstracted from it* [12]. Let us examine Searle's notion of proposition which is of this type.

> "Imagine a speaker and a hearer and suppose that in appropriate circumstances the speaker utters one of the following sentences:
>
> 1. Sam smokes habitually.
> 2. Does Sam smoke habitually?
> 3. Sam, smoke habitually!
> 4. Would that Sam smoked habitually.
>
> ...anyone who utters one of these can be said to have uttered a sentence formed of words in the English language. But clearly this is only the beginning of a description, for the speaker in uttering one of these is characteristically saying something and not merely mouthing words. In uttering 1 a speaker is making (what philosophers call) an assertion, in 2 asking a question, in 3 giving an order, and in 4 (a somewhat archaic form) expressing a wish or desire. And in the performance of each of these four different acts the speaker performs certain other acts which are common to all four: in uttering any of these the speaker *refers to* or mentions or designates a certain object Sam, and he predicates the expression "smokes habitually" (or one of its inflections) of the object referred to. Thus we shall say that in

[12] Cf. John R. Searle, "Austin on Locutionary and Illocutionary Acts", in J.F. Rosenberg and C.Travis, eds., op.cit., E.Stenius, "Mood and Language Game", Synthese, Vol.17, 1967, Lennart Aqvist, "Semantic and Pragmatic Characterizability of Linguistic Usage", Synthese, Vol.17, 1967.

> the utterance of all four the reference and predication are the same, though in each case the same reference and predication occur as part of a complete speech act which is different from any of the other three. We thus detach the notions of referring and predicating from the notions of such complete speech acts as asserting, questioning, commanding, etc., and the justification for this separation lies in the fact that the same reference and predication can occur in the performance of different complete speech acts. Austin baptized these complete speech acts with the name "illocutionary acts"...
>
> Whenever two illocutionary acts contain the same reference and predication, provided that the meaning of the referring expression is the same, I shall say the same proposition is expressed. Thus, in the utterances 1-5, the same proposition is expressed. And similarly in the utterances of:
>
> 6. If Sam smokes habitually, he will not live long.
> 7. The proposition that Sam smokes habitually is uninteresting.
>
> the same proposition is expressed as in 1-5, though in both 6 and 7 the proposition occurs as part of another proposition. Thus *a proposition is to be sharply distinguished from an assertion or statement of it* , since in utterances of 1-7 the same proposition occurs, but only in 1 and 5 is it asserted. Stating and asserting are acts, but propositions are not acts. A proposition is what is asserted in the act of asserting, what is stated in the act of stating. The same point in a different way: an assertion is a (very special kind of) commitment to the truth of a proposition. ...
>
> I might summarize this part of my set of distinctions by saying that I am distinguishing between the illocutionary act and the propositional content of the illocutionary act." [13]

Searle proposes an analysis of uttered sentence tokens which would distinguish between (a) an illocutionary force indicator, representing those aspects of the uttered sentence relevant to the determination of its illocutionary force (such as the presence of certain performative verbs, word order, stress, the mood of the verb, and so on), and (b) a propositional indicator, representing that aspect of the utterance which is neutral to illocutionary force, viz., the proposition expressed. This propositional indicator might also be called a "sentence radical".

We have a wealth of different notions of 'proposition' to choose between. Rather than debate their respective values, we must now ask what notion is closest to that which the team would like to investigate.

[13] John R. Searle, "Speech Acts", Cambridge University Press, 1969, pgs. 22-30.

I.2.2 THE TEAM'S NOTION OF (REPEATED) PROPOSITIONAL REFERENCE

The dialogue analysis team is at present seeking to formulate a notion of proposition congruent with its research interests, and to employ that notion in selecting a set of phenomena which it feels it is presently ready to examine.

I.2.2.1 The team's notion of proposition. No description exists in print as yet of what the team's concept of proposition or of propositional reference might be (propositional, as opposed to text reference, was not mentioned in the "Observation Methods" report). The following remarks are therefore based on discussions I have had with members of the team, particularly with Jim Levin.

It would seem that the team is moving towards a notion of proposition which is much more akin to that of Searle and of some linguists than to that of the 'logicians'. The team is employing a notion of proposition as a theoretical notion employed in the context of the modeling of dialogue by means of semantic nets. This notion is to be understood by reference to a certain form of representation of utterances in a dialogue. In order to see this more clearly, consider the utterances in (1)-(3) below, accompanied by one form of representation which the team might employ. (In these examples, we suppose that Bill is addressing his utterances to John.)

(1) Bill: You will shut the door.

```
(SAY
  (BILL
   TIME-1
   JOHN
   (SHUT
        (JOHN
         DOOR-1
         TIME-2))))
```

(2) Bill: Will you shut the door?

```
(ASK
  (BILL
   TIME-1
   JOHN
   (SHUT
        (JOHN
         DOOR-1
         TIME-2))))
```

(3) Bill: John, shut the door!

```
(ORDER
  (BILL
   TIME-1
   JOHN
   (SHUT
        (JOHN
         DOOR-1
         TIME-2))))
```

The three representations have a common element, namely, (SHUT (JOHN DOOR-1 TIME-2)). This common element is very similar to what Searle calls a proposition: it is a predication abstracted from different illocutionary acts (and not only from statements or assertions). This common element is part of what is a proposition expressed by a declarative sentence, e.g., in the case of (1), the proposition expressed by a sentence such as "Bill asked John at time t1 to shut the door at time t2". This matrix sentence or proposition contains an illocutionary verb ("say", "ask", "order", etc.) and thus conveys not only the proposition expressed by the common element, but also the illocutionary force with which that common element was produced. Notice, however, that the representation is one in which indexicals are filled in, and in which the referents of noun-phrases in the utterances are unambiguously specified (thus DOOR-1 is a GENSYM); in this respect the representation is akin to the logicians' representations of statements.

Given this form of representation, one might define proposition in one or more of several ways. One might reserve the term for the representation of an act of uttering a sentence taken in its entirety. One might term proposition any complete representational unit, i.e. any verb and its arguments; a proposition in this sense would include, of course, propositions in the first sense above. Or one might exclude the matrix representation, and only term its components propositions; these propositions would then correspond to (nominalized) sentential clauses in the utterances. To illustrate these possibilities, let us consider the utterance and its representation below:

Bill: I am sick, and I believe that I am going to faint.

```
(SAY
  (BILL
   TIME-1
   JOHN
   (AND
        (IS BILL SICK TIME-1)
        (BELIEVE
                (BILL
                 TIME-1
                 (FAINT
                        (BILL
                         TIME-2]
```

The entire representation, [SAY ...] is a proposition in the first sense. Propositions in the second sense include (a) [SAY...], (b) [AND...], (c) [IS...], (d) [BELIEVE...] and (e) [FAINT...]. Propositions in the third sense include only (b)-(e), and not (a).

At the present stage of discussion, I can only throw out these alternatives for the sake of debate.

I.2.2.2 The team currently has a clearer notion of just what it is that they wish to explore under the heading of the term "repeated propositional reference" than they do of what they wish to define their notion of proposition as. So let us now turn to a consideration of some dialogue phenomena which the team (a) feels are instances of repeated propositional reference, and (b) are interested in investigating.

The following is a series of constructed examples of dialogue excerpts which the team would feel involve instances of repeated propositional reference phenomena of a type currently worthy of analysis.

THE NOTIONS OF TEXT AND REPEATED PROPOSITIONAL REFERENCE

(1) A: Mary is sick.
B: That's unfortunate.

(2) A: I have no money. Five members of my family are deathly ill. I can't sleep at night because I have to keep fighting off the blood-thirsty rats in my vermin-infested apartment block.
B: Your story is not so different from the stories most other people in the neighborhood could tell.

(3) A: My daughter is expecting a child.
B: Yes, but it's a great secret; don't tell anyone else about it.

(4) A: John said that Mary is sick.
B: Paul told me that piece of bad news yesterday.

(5) A: John said that Mary is sick.
B: If what he said is true, we can't have our picnic.

(6) A: John believes that Mary is sick.
B: Yes, but Paul doubts it.

(7) A: John believes that Mary is sick.
B: What he believes is true, unfortunately.

When one first rapidly glances at the above examples, one feels that they are similar in that in each dialogue some pronoun or noun phrase is used to talk about something which has previously been talked about by the use of some sentence(s) or nominalized sentential phrase. One also feels that this similarity can only be specified in very vague terms, as was just done. And a closer look at the examples shows why: there is an extraordinary variety of things going on. In (1), the first sentence is used to describe a state of affairs, which is then referred to and commented on by the second utterance. In (2), the first turn involves a description of a state of affairs, and the second utterance comments not so much on that state of affairs as on the description which was made of it (even though one would clearly hesitate to say that a text reference was being made). In (3), the first turn involves the imparting of a piece of information which is a description of a state of affairs, but which is commented upon as a piece of information in the second turn (the information is true, but its a secret - i.e. it is a piece of information which has not been imparted to many people). In (4), indirectly quoted speech is reported, and then is said to be identical with some other reported speech; (5) again involves reported speech, but the statement which was said to be made is then treated as a proposition in a "transparent" context. In (6) and (7), an intentional object - a belief - is talked about, but in (6) it is talked about within an opaque context by both participants, whereas in (7) a belief is talked about once in an opaque context and again in a transparent context.

This heterogeneity is not such as to indicate that the PREF phenomena which the team currently wishes to study are completely ill-defined, however. First of all, the team

is well-united on the decision that they do not wish to study PREF phenomena which involve relationships between two synonymous sentences used in a dialogue; nor do they wish to study under the heading of PREF the relationship which holds between two utterances in a dialogue such that the participants who uttered them thereby "said the same thing" or made the same statement. Thus the following pairs of sentences, though they are instances of PREF, are not of current interest.

A: Mary is sick.
B: Mary is ill.

A: You are hot.
B: Yes, I am hot.

Secondly, there would seem to be some consensus on not including in the study of PREF an investigation of the relationship of co-referentiality between noun-phrases which denote what we would ordinarily regard as propositions or statements. NP-NP co-reference is thus outside of the scope of current PREF analysis, and such dialogues as those below are not to be studied under that heading:

A: John's assertion is simply not true!
B: I can't see why not. His claim seems well supported by all the available evidence.

A: De Morgan's law is a very important one.
B: Well, his theorem has certainly been useful.

Thirdly, there are certain referential phenomena which are, intuitively, quite distinct from what we are groping at above. One is reference to physical objects; another is reference to text per se. A third phenomena which is distinct is reference to actions, as in (8) and (9) below:

(8) A: John went fishing yesterday.
B: Mary did so too.

(9) A: Sky-diving without any training is exhilarating.
B: It's also foolish and suicidal. [14]

The distinction between propositional reference and action reference is often intuitively clear, but it is very hard to formulate. It cannot be pinned down in syntactic terms alone. One is tempted to say that neither the "it" nor the "so" in the above examples are propositional references because they cannot be analysed as standing for sentences, nominalized or no. The second turn in (8) could be rendered as "Mary did go fishing yesterday too" but not as "Mary did John went fishing yesterday too". However, the second turn of (9) could be rendered as "For people to go skydiving without any training is also foolish and suicidal". And the noun phrases or pronouns involved in PREF often

[14] An interesting example of reference to an action which can occur in dialogue is reference to a performative act previously performed by a participant, as in:
A: You're a bastard!
B: That's un-called for!

cannot simply be replaced by sentences: consider "your story" in example (2).

Given the above, we will say that repeated propositional reference of the type the team is currently interested in studying occurs in a dialogue when

(1) a sentential or multi-sentential utterance in the dialogue is subsequently mentioned by means of a pronoun or noun-phrase in such a manner that what is being mentioned is not the utterance considered purely as text (i.e. as an uninterpreted phonetic or graphic token, or as a member of some class of tokens).

(2) Some sentence uttered in the dialogue contains a nominalized sentence which refers to a state of affairs, a statement, a reported utterance or an intentional object, and some subsequent pronoun or noun-phrase is co-referential with that nominalized sentence.

Syntactically speaking, then, PREF involves only sentences and nominalized sentences on the one hand, and noun-phrases and pronouns on the other (see diagram overleaf).

The above delimitation of PREF is highly criticisable: it is a description by exclusion (cf (1)), and contains highly problematic terms (e.g. "state of affairs"). However, it is the least bad proposal I can come up with.

There is a terminological problem which remains to be dealt with. It is unsatisfactory to use the term "propositional reference", to describe a phenomena which covers some phenomena (those described in (1)) which do not involve reference in the usual sense at all. I shall continue using the term PREF, but only until a better term is found.

SECTION II:
DETECTION PROCEDURES
FOR
TEXT AND REPEATED PROPOSITIONAL
REFERENCE

In this section, we will discuss constraints on TREF and PREF, as described above, which might be of some assistance in detecting their occurence by analysis of dialogue transcripts.

These constraints will be induced from an examination of dialogue samples in which TREF and PREF appear to occur. The examples will be of three types: some will be examples of real dialogue, some will be examples of dialogues drawn from literary works, such as plays or short stories, and some will be examples which have been made up (constructed) to illustrate a point or a difficulty. Each example provided will be marked by an R, an L or a C, according to its source.

The first observation which one makes when one considers actual dialogue is that TREF is much less frequent than PREF. Since TREF is more of an exceptional phenomena, we will examine it first, with the hope that occurences of it may be signaled more explicitly than are occurences of PREF.

II.1 THE DETECTION OF TEXT REFERENCE

As we have defined or described it above, TREF always involves the use of at least one expression to refer to text. A text-referring expression (TREXPR) [15] may be an ordinary noun-phrase ("his name", "what John wrote on the board"), or may be an expression which refers to text by either exhibiting or replicating it. So if we wish to detect instances of TREF, we must (a) find ways of determining whether an expression is being used as a TREXPR.

But once we have determined that a given expression is a TREXPR, we are not thereby assured that we have before us an instance of TREF as we have conceived of it. For a particular TREXPR, say E1, to be involved in TREF, one of the following two cases must obtain. (i) There is another token, say E, used in the dialogue, which is not a TREXPR, and which is either identical with or a member of the referent of E1. (ii) There is another TREXPR, say E2, used elsewhere in the dialogue, whose referent is either identical with, a subset of, or a member of the referent of E1. Thus, once we have detected the presence of a TREXPR in a dialogue (a), we must (b) compare its referent to other non-TREXPRs in the dialogue and to the referents of other TREXPRS in the dialogue, if there are such, and (c) decide on the basis of this comparison whether a TREF is occuring, and if so, what other expression is involved.

The above suggests an outline of a procedure for detecting TREF:

1. Find all occurences of TREXPRS.
2. For each TREXPR:

[15] For the remainder of section II, unless otherwise specified, we will use the term "expression" to mean expression-token.

2.1 Draw up a list of candidates for comparison with the TREXPR. This list will include
- other TREXPRs, if present
- "suitable" tokens, not TREXPRs, which are used in the dialogue.

2.2 Compare the TREXPR with the candidates.
- If the candidate is a TREXPR, determine whether its referent is identical with, a subset of, or a member of, the referent of the TREXPR under study.
- If the candidate is a non-TREXPR, determine whether it is identical with or a member of the referent of the TREXPR under study.

If no comparisons suceed, conclude that no TREF is occuring.

If only one comparison succeeds, return the pair of successfully compared expressions as the TREF which is occuring.

If several comparisons succeed, continue.

2.3 Apply some evaluative criterion (or criteria) to the pairs of expressions which have been successfully compared. If one pair is clearly a "best" match, return it as the TREF; if several pairs are almost equally "good", return the TREXPR under study along with all other members of these pairs as the (multiple) TREF which is occuring.

With this vaguely defined procedure in mind, let us turn to a consideration of a corpus of examples of text reference (listed overleaf).

The corpus consists of dialogue examples which are either real or literary (fictional). The real examples are either examples of written dialogue or of oral dialogues which were subsequently transcribed; the literary examples were written. We will give less weight to confirmation of procedures by transcribed examples when those procedures rely upon orthographic cues (such as the presence of quotation marks or capitalization). The reason for this is that orthographic cues are furnished not by the original participants in the dialogues but by the transcriber, and are the result of a decision by the transcriber that some expression was being used as a TREXPR; to rely upon orthographic cues in such cases is a 'cop-out' from the point of view of someone who wishes to to specify computer programs which will detect TREF independently of human judgments.

We will consider examples in turn, proposing subprocedures or criteria for each which will then be applied to subsequent examples. We will at first only deal with noun phrases, which present - on the whole - fewer problems for analysis, and then go on to deal with pronouns.

Let us begin with example (1) below. (In all of our examples we will italicise the expressions which we feel intuitively are involved in a TREF, and number them for ease of subsequent discussion.

L: Yes, I linked to PARC-MAXC and transferred a file - but couldn't run it at PARC-MAXC - *IDFAS?* (1)

O: What's *IDFAS?* (2)

L: Sorry, mistyped - ideas?

(REAL - WRITTEN - SOURCE: OC32.PROTOCOL)

We intuitively perceive the second occurence of "IDFAS", (2), as a TREXPR. This suggests that we regard expressions which are not in our lexicon either as words of English or as names as TREXPRs; but this principle is wrong, because it would lead us to mark the first occurence of IDFAS, (1), as a TREXPR also, wheras we clearly percieve it to be a simple expression (albeit a mistyped version of an expression) which L uses. Similarly, simple capitalization is not a cue, for both occurences of "IDFAS" are capitalized. What seems to indicate to us that (2) is a TREXPR is not only that it is an unrecognized symbol, but that it is the subject of a question. So this leads to the formulation of a principle of TREXPR detection (TD):

(TD1.1)

If an expression is not in the lexicon, then
if it is the subject of a question,
it is a TREXPR.

This principle in the case of example (1) allows us to conclude that there is one and only one TREXPR. What are candidates for comparison with it? Intuitively, we perceive that there is only one non-TREXPR candidate: the first occurence of "IDFAS", (1), which is a token of the type of the TREXPR (2) (this is an instance of TREF by replication). This leads us to formulate the following principle of candidate selection (CS):

(CS1.1)

If there is a non-TREXPR expression in the dialogue
which is orthographically identical with the
TREXPR under study, then it is a candidate.

Since there is only one candidate, our procedure returns the two occurences of "IDFAS" as a TREF.

Let us now consider example (2) below:

L: Guess what ... that didn't work either. It took *"TTY"* (1) to be a filename.

O: Did you say *TTY* (2) or *TTY:* (3) ?

L: Just *TTY*. (4)

O: If you append the colon, then it will be recognised as a device designator instead of file name (I hope).

What are the TREXPRs here? Principle TD1.1 picks out expressions (2) and (3). But intuitively we know that the quotation-expression (1) is a TREXPR, because it is explicitly quoted. Wo we have a new TD principle:

(TD2.1)

If an expression is enclosed in quotation-marks,
then the quotation-expression (i.e. the
quotation-marks plus what they enclose)
is a TREXPR.

We still do not have sufficient TD principles, however, for intuitively we perceive the occurence of "TTY" in L's second turn (4) to be a TREXPR also, and neither TD1 nor TD2 would mark it as such. I feel that I recognise that token to be a TREXPR for two reasons: (a) it is an expression which is not in the lexicon, and (b) it is said to have been uttered by someone, viz. L (L's utterance is elliptical, but expandable into "I just said TTY.") In all of the constructed sentences below, one would detect a TREXPR:

John said/is saying/says/will say blurpagg.
Mary shouted/is shouting/shouts/will shout ARRGGHH.
I wrote/am writing/write/will write Xuytmon.

Note that linguistic-reception verbs have the same effect as these linguistic-production verbs ("hear" as well as "say"). So we modify TD1 as follows:

(TD1.2)

If an expression is not in the lexicon, then
if it is the subject of a question,
it is a TREXPR;
if it is the object of a verb of linguistic
production or reception,
then it is a TREXPR.

So now our principles allow us to recognize four noun-phrase TREXPRS: (1) in turn 1, (2) and (3) in turn 2, and (4) in turn 3. Applying CS1.1, we find that there are no non-TREXPR candidates. So we are left to take each of the four TREXPRs in turn and compare them with the three others.

We percieve (4) to be co-referential with both of the other occurences of "TTY", and we formulate the following principle of co-referentiality of TREXPRS (CR) to account for that fact:

(CR1.1)

If there are two TREXPRs, E1 and E2, and one of
them, say E1, is a quotation-expression, then
if E2 is orthographically identical with
the quotation-content (i.e. the string
between the quotation-marks) of E1,
then E1 and E2 are co-referential.

We perceive the first and second occurences of "TTY", (2) and (4), to be co-referential, which can be expiained as foiiows:

(CR2.1)

If there are two TREXPRs, E1 and E2, and they are
orthographically identical,
then they are co-referential.

We also perceive each of the occurences of "TTY" to be co-referential with each other and with ""TTY"". This can be seen as a logical consequence of the fact that co-referentiality is an equivalence relation. One would be wrong in so concluding, however, for one must

remember that we are counting as co-referentiality in matters of TREXPRs (a) identity of referent (b) subset-superset relations of referents, and (c) member-set relations of referents. And of course, if x is a subset/member of a set z, and y is also a subset/member of z, it does not follow that x and y are either identical or bear subset-superset/member-set relations to each other. Nevertheless, I will temporarily adopt the following heuristic, with full knowledge that it is false, but with an eye to the fact that its falsity may not be revealed in most dialogues.

(CR3.1)

For any three TREXPRs, E1, E2, and E3, if E1 and E2 are both co-referential with E3, then E1 and E2 are co-referential with each other.

Lastly, we do not perceive (3) to be co-referential with any other noun-phrase TREXPR. This follows from the above CR principles.

In example (2) above, we relied upon the presence of verbs of linguistic production and reception to detect TREF. Certain nouns may also signal possible TREFs. Consider example (3) below:

O: Yep, and I will do my best to help. What is it you want? [name 1] Go ahead.

L: I would like to unarchive tapes 1120 and 1121 programs are called */name 2/* (1) ...~~~~~~~(interrupt here)~~~~~~~~

.... We would like to unarchive these.

O: in?

L: We are in directory [name 3] but the tapes were archived from the [name 4] directory. Go ahead.

Ok, but you will have to give *me those names* (2) again...

(REAL - WRITTEN - SOURCE: OC636.PROTOCOL)

We perceive expression (2) to be a TREXPR because of the meaning of the word "name": a name is a type of sound or inscription which we use to refer to individuals. So we adopt the following TD heuristic:

(TD2.1)

If an expression is a member of the set of conventional TREXPRS (C-TREXPRS), then it is a TREXPR.

(TD3.1)

The set of C-TREXPRS is <name(s)>.

We also perceive the expression (1) to be a TREXPR. The reason for this at first seems that it is the indirect object of the verb "call", which suggests the following general principle:

(TD4.1)

If any expression is the indirect object of the verb "call", then it is a TREXPR.

The principles TD3 and TD4 seem to be borne out by the following example (4):

A: Our president calls us "*Ambassadors of Friendship*". (1)

B: *Beautiful motto.* I winder if you know a conductor by the name of *George Whipple* (3) ?

A: George Whipple? No. I knew a *George Calloway* (2).

B: This is George Whipple.

A: I don't recall *the name* (4).

(LITERARY - WRITTEN - SOURCE: THE TRAVELLOR)

We can see that (3) and (4) are TREXPRs using TD4.1. (4), of course, requires semantic processing to establish that "by the name of X" introduces a textual object, X, which is a name.

We have been concentrating so far on noun-phrases, trying to decide whether they are TREXPRs, and what other noun-phrases or expressions TREXPR noun-phrases are associated with in occurences of TREF. Given the presence of "this" in B's second turn in example (4), we should begin to consider examples of TREF which involve pronouns. Here it is useful to consider certain general rules for finding the antecedents of pronouns which are useful in cases of repeated reference not only to text, but also to actions, objects and propositions. We shall only be concerned here with general rules for the determination of the antecedents of the pronouns "it", "that" and "this", given the fact that text is never referred to by the pronouns "he" or "she", because of gender considerations.

(GR1.1) IF A PRONOUN = "IT", "THAT" OR "THIS":

(1) If there is another previous pronoun of the same type in the same sentence, then that pronoun is a candidate of priority 1 for co-referentiality;

(2) If there is a pronoun in the nth preceding sntence (where n is less than some parameter I), then that pronoun is a candidate of priority n+1 for co-referentiality.

(3) If no pronominal candidates are found, then preceding noun-phrases are candidates.

(a) Noun phrases within same sentence have a higher priority than noun phrases in preceding sentences.

(b) Noun phrases in nth preceding sentence have lower

priority than noun-phrases in mth preceding sentence, where n and m are both less than some parameter I and n<m.

(c) Within a sentence, noun phrases in subject position have a higher priority than noun phrases in object position; noun phrases in object position in turn have a higher priority than non phrases in prepositional phrases.

(d) Within a sentence, noun-phrases within a main clause have a higher priority than noun-phrases within a subordinate clause.

(e) Within a sentence, focused noun-phrases have a higher priority than noun-phrases which are not focused. (Focused noun-phrases are those moved to the front of sentences by such transformations as extraposition or tough-movement.)

(GR2.1)

Candidates for co-referentiality are to be selected by comparing the following features or predicates of the pronoun with those of the candidate: (a) sex and number (b) case or type constraints. If such considerations fail to select one candidate, rely on more specific plausibility considerations. [16]

Let us see how these general rules, in conjunction with the rules we have specified so far, allow us to detect TREF. Consider first the following example (5):

A: You've surely heard me speak of *Eugene Tesh* (1)!

B: I can't say that I have.

A: Well, *his name* (2) is always in the newspapers; he's a dramatic star. Everyone I know would recognize *it* (3).

B: I am not familiar with *the names of dramatic stars* (4). I have never seen *it* (5) before.

(CONSTRUCTED)

[16] These general rules were suggested by Jim Levin, on the basis of his own work and the work of others in the AI field. I do not claim that he would endorse the form that I have given them here.

According to our present rules, "Eugene Tesh" is marked as a TREXPR because it is not in the lexicon. When we get to expression (2), which is marked as a TREXPR by rules TD2 and TD3, (1) is the only candidate for partnership in a TREXPR. Now we understand that when reference to an individual is made by the introduction of his/her name, two pieces of information are imparted: (a) that there is an individual names X, and that there is a textual object, viz. X itself, which bears the relation name-of to that individual. And in this particular case, we understand that (2) is a reference to the name "Eugene Tesh", and not to the individual named Eugene Tesh. We might capture this by the following crude rule:

(CR4.1)

If there are two TREXPRs, one of which is or contains the word "name" or "names", and the other one of which has been interpreted as referring to an individual by name, then they are co-referential.

We are supposing that the phrase "interpreted as referring to an individual by name" has some meaning in terms of syntactic and semantic programs yet to be specified. When we come to the pronoun (3), the general rules comes into play. According to these rules, the first expressions which would be considered as candidates for antecedents of (3) would be noun-phrases, since there are no "it"s, "that"s or "this"s in the preceding sentences. Noun-phrases having higher priority than (2) would be rejected because of sex or number constrain's (e.g. "a dramatic star"). And (2) would be selected. When we come to (4), CR4 would mark it as co-referential with (1), and CR3 would mark (1), (2), (3) and (4) as co-referential. Lastly, the general rules would mark (2) as the antecedent of (5), and CR3 would again chain (1)-(5) together as co-referential.

Another example in which our present rules would give us a satisfactory result would be the following example (6):

A: Fine! The other one is a *L.V.N.* (1) down at Permanente on Sunset Blvd.

B: *L.V.N.* (2) ... *That's* (3) lanky vertiginous nurse?

A: Right.

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

Our rules would mark (1) and (2) as TREXPRs, by TD1. We perceive (1) and (2) to be co-referential, by:

(CR5.1)

If two TREXPRs are not in the lexicon, then
if they are orthographically identical,
then they are co-referential.

When we come to the pronominal expression (3), the general rules would mark (2) as being the first noun-phrase acceptable candidate for being (3)'s antecedent. Note that this conclusion is very fragile. If B had not repeated "L.V.N.", the general rules would have selected "Sunset" as (3)'s antecedent.

A furthur example showing the additional development of the general rules is the following, example (7):

A: I wonder if you know a conductor by the name of *George Whipple* (1) ?

B: George Whipple? No. I knew a *George Galloway* (2).

A: This is George Whipple.

B: I don't recall *the name* (3).

(LITERARY - WRITTEN - SOURCE: THE TRAVELLOR)

Our rules mark (1) as providing a TREXPR; as above, we shall assume that the processor will, in the interpretation of A's first turn, (i) introduce an individual named George Whipple and (ii) introduce a textual object, "George Whipple", which bears the relation name-of to that individual. Similarly for (2). But now a difficulty becomes apparent. The pronoun "this" in "This is George Whipple" clearly refers to the individual referred to in the first turn, and not to his name. Our general rules would probably bind "this" to "George Galloway", and if they did not (on general grounds of the implausibility of the resultant conclusion that George Galloway is George Whipple), they would bind "this" to the name, "George Whipple".

The above gives the flavor of the process of gradual development which might well lead to some acceptable rules for detecting TREF. Obviously, what has been said above only represents the beginning of such a process. I would like to end this section by mentioning some of the difficulties which the construction of TREF rules will undoubtedly encounter.

First of all, there are instances in which people make spurious use of orthographic cues of TREF. Consider example (8):

A: Hello. Got a couple of questions about "runoff" to on-line. Go ahead.

B: Okay. I've got a manual here, and although I don't know too much about it, we'll see what I can find. Hold a sec ... O.K. Shoot.

A: I have a rather old manual and I am trying to get runoff to print to my TTY on line...

(REAL - WRITTEN - SOURCE: OC370.PROTOCOL)

The use of quotation-marks around "runoff" in the first turn might well be regarded as a simple mistake (though there is an alternative interpretation, to which we will return below). If it is so regarded, the question arises: should one build rules that will test for and eliminate mistakes?

Secondly, there is the problem of TREXPRs which contain variables. An example would be "When making a new connection type ATT LINKER (PASSWORD), job number". Here only the "ATT" is a replication of its referent; "LINKER" stands in for a variable string, or a variable class of strings. Logicians have wrestled with the problem of variables within quotation contexts because they needed to be able to have variables ranging over text when constructing truth definitions; there are considerable problems here.

Thirdly, there is the problem of the use of quotation-marks in irony, the use of so-called snigger quotes. Jim Levin has suggested what is undoubtedly the right approach to such problems: regard quotation-marks in general as a signal that some peculiarity in processing the quoted words is required. The most frequent peculiarity thus signaled is that the words themselves should be retained, but other peculiarities should be allowed for. For example, quotation-marks may signal that a word is being used in a sense very different (perhaps opposite) from that in which it is commonly used, or that some presupposition of the use of that word is not obtaining. To return to example (8) above, for instance, it is possible that the user was quoting "runoff" to show that though the term usually is supposed to denote a program that runs off formatted copies, he is unwilling to use it in that way, because his experience leads him to believe that the program in question obstinately refuses to run off copies!

Despite these difficulties, I believe that one might well be able to devise a set of heuristics which would detect TREF correctly in a satisfactory number of cases. TREF is on the whole much less intractable than PREF, to which we now turn.

II.2 THE DETECTION OF REPEATED PROPOSITIONAL REFERENCE

Our general approach to the detection of repeated propositional reference will be largely similar to our approach to the detection of TREF. That is, we will first look for certain proposition-referring expressions, PREXPRs, and then, once PREXPR(s) have been found, look for other expressions which may be involved with those PREXPR(s) in repeated propositional reference. Lastly, once we have a PREXPR and a list of "candidates", we will select candidates by some criteria.

There will be important differences, however. PREF involves (a) a noun-phrase or pronoun on the one hand, and (b) a sentence or nominalised sentence, on the other. Now the first conclusion one reaches when one considers instances of PREF is that *just about any sentence or nominalized sentence is capable of participating in a PREF*. As a result, it is not functional to first pick a sentence or nominalized sentence and then look for a noun-phrase or pronominal candidate - there would simply be too much useless processing involved in such a procedure. Rather, one must first find a noun-phrase or pronoun which, because of its meaning, grammatical position or features (i.e. because of what is predicated of it) is suceptible of being involved in a PREF, and then look for sentence or nominalized sentence candidates.

Below, we will first analyse examples of PREF which involve noun-phrases, and then look at examples of PREF which involve pronouns.

First, let us consider example (1) below:

A: You know, I just... *The second question* (1) would be: *why wasn't this done before I went through all these bone scans, thyroid scans, and you know...* (2)

B: Well, that's a...

A: I'll die of radioactivity.

B: Yes, that's *the logical question* (3) and...

(REAL - ORAL - SOURCE: MEDICAL-CENTERED.PROTOCOL)

We see intuitively that "question" is a noun-phrase which refers to what we would call a proposition, so that both (1) and (3) are noun-phrases which might be involved in a PREF. Other noun phrases are similar to "question" in this respect, for example "statement", "request", "order", "demand", "query" and so on. This suggests the following principle for detecting PREXPRs (PD):

(PD1.1)

If an expression is a member of the set of conventional proposition-referring expressions (CPREXPRs), then it is a PREXPR.

(PD2.1)

The set of CPREXPRs is <question(s), request(s), order(s), demand(s), assertion(s), claim(s), query(ies)....>

We also see that (2) is co-referential with both (1) and (3), because it is an expression of the interrogative form. What we are relying on here is a syntactic requirement which derives from the meaning of the noun "question". Certain other CPREXPRS impose syntactic requirements upon candidates, for example, "advice" in the following example (2):

A: And now I shall give you *an extra piece of advice* (1). *Stop disgracing your daughter with your company on the streets - and, above all, at the theatre...* (2) or she will soon have every door to advancement shut to her!

(LITERARY - WRITTEN - SOURCE: MOTHERLOVE)

"Advice" is a CPREXPR, and we know that (2) is a candidate for PREF because it is in the imperative mood. A third example of a dialogue in which syntactic cues deriving from the meaning of a CPREXPR are used in determining PREF is (3) below:

A: General, I only want to keep one little private letter. Only one. *Let me have it.* (1)

B: Is that *a reasonable demand* (2), madam?

(LITERARY - WRITTEN - SOURCE: THE MAN OF DESTINY)

Here again, we see that (2) is co-referential with (1) because demands are (usually - I am being sloppy here of course) expressed by means of imperative sentences. One may therefore adopt the following heuristic:

(TD3.1)

Given a CPREXPR, if that CPREXPR denotes a type of proposition which is usually expressed by a sentence of a certain grammatical mood (declarative, interrogative, imperative, horartory, etc.), then any nearby sentence or clause which is of that mood is to be considered as a PREXPR cadidate.

We have included the phrase "sentential clause" in TD3.1 because we want to account for cases like "Sam is curious to know whether or not the Socialists will take over the Portugese government. - That's a good question". In such cases one finds that embedded questions participate in PREF.

Many instances of PREF involving noun-phrases require a pretty complete understanding of the meaning of the noun-phrase in question in order to select candidates. I will give three examples below. First, example (4):

A: Just a passing *comment* (1), Joe. We're having lunch right now, and I just made myself a hotdog sandwich with catsup. *Very tasty and almost unheard of in the old days.* (2)

(REAL - ORAL - SOURCE: CORRECTION-ACTIONS-CORPUS.PROTOCOL)

(I will suppose that the ellipsis in (2) has been filled out.) In deciding that (2) rather than the preceding sentence is a candidate for PREF involvement with the CPREXPR (1), we rely upon our understanding of what a comment is, i.e. an observation or remark expressing an opinion or attitude. Similarly, in example (5) below:

A: ... And, Mary, I can tell you *a secret.* (1) It's still *a great secret* (2), mind! *They're expecting a grandchild.* (3) Isn't that *good news* (4)?

(LITERARY - WRITTEN - SOURCE: THE LONG CHRISTMAS DINNER)

our selection for candidates for (1) and (2) and for (4) depends (a) on our knowledge of what a secret and what news is, and (b) on what kind of information would probably count as a secret or as news to the participants given the situation and the participants' knowledge. Lastly, consider example (6):

A: So, anyway when we got there *the funniest thing* (1) happened. *They sat down and they passed out these little booklets* (2), because we went to their suite.

B: Uh-huh.

A: *And, they started preaching about their religion the whole three hours* (2) and we were just crawling the walls to get out.

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

This is a complex example. First, we recognize (1) as being a PREXPR, not because it is itself a CPREXPR, but because it is the subject of a verb which takes as subjects nouns which refer to events, e.g. the verb "happen". This leads to a new principle:

(PD3.1)

If an expression is the subject of an event-verb,
then it is a PREXPR.

(PD4.1)

The set of event-verbs is <happen, occur, ...>

We also know that candidates for involvement in PREF with event-nouns must be declarative sentences which describe events or states of affairs, e.g. they cannot express general laws. I hesitate to make a rule of candidate selection out of this intuition, however, because I can think of no operationalizable way of detecting when a declarative sentence describes an event or state of affairs. (This requires furthur work, to say the least!) But let us return to our previous comment about the need to understand the meaning of PREXPR-nouns in order to select candidates. We intuit that it is not a strange or funny thing that the speaker and the group of people the speaker was with "started

crawling the walls to get out", given the circumstances. It is this complex understanding which allows us to decide that the last part of the third turn is not part of the PREF. We also intuit that the last clause of the first turn is not part of the PREF, and we do so not only because we understand that, in the circumstances described, it was not a strange thing for the speaker to go to their suite, but also because of our understanding of the semantic function of the clause "because we went to their suite". There are two different functions of clauses prefaced by the word "because", illustrated by the two following sentences:

(a) John is not coming to the meeting tonight,
because he is sick.

(b) John is not coming to the meeting tonight,
because he just phoned me from Australia.

"Because" can either be used to talk about causes, as in (a), or to introduce considerations which either logically or plausibly justify making a certain statement, as in (b). In the example above, "because" is being used to explain or justify the making of a certain description, and is thus not part of that description itself.

Let us now turn to an examination of some cases of PREF which involve pronouns.

Some general heuristics may be laid down at the outset. The first concerns the distinction between pronouns which refer to actions and pronouns which refer to propositions. Consider the following example (7):

A: And, for your information, Jack, I'm just going *to tear into some beef and gravy and other assorted goodies.* (1)

B: I presume that you're doing *this* (2) with the full permission and - of the commander.

(REAL - ORAL - SOURCE: APOLLO-13/PAGE379.PROTOCOL)

Here we intuit that the pronoun (2) is involved in what might be called a repeated reference to an action. The principle clue is that (2) is the object of the pro-verb "do". This clue, in so far as I have been able to ascertain, is a frequent and reliable one. It is not the only clue, however, as the example (8) shows:

A: We have decided to use a canister and, you know that the liquid-cooled garment has a bag around it that we think we can use too, or that we know we can use. We've tried *it* (1).

(REAL - ORAL - SOURCE: APOLLO-13/PAGE379.PROTOCOL)

Here we intuit that (1) is involved in a repeated action reference, although just what action is involved is rather unclear. The clue here is that the pronoun (1) is the object of the verb "try". So we can formulate the following rather solid heuristic:

(PD5.1)

If a pronominal expression is the object of either of

the verbs "do" or "try", then it is not a PREXPR.

Secondly, we know that there are certain predicates which mark pronouns as PREXPRs. Some of these can be grouped into two classes: the class of what can be loosely called "logical" predicates, such as "x is true/false" , "x is possible/probable", "x is inconsistent" or "x implies y", and what can be called intentional predicates, such as "x knows/believes y".

(PD6.1)

If an expression is such that some logical predicate is attributed to its referent, then it is a PREXPR. Logical predicates include the adjectives "true", "false", "probable", "possible", and the verbs "imply", "entail".

(PD7.1)

If an expression is the object of an intentional verb, then it is a PREXPR. Intentional verbs are a class of verbs which includes "know", "believe", "remember", "wants".

As soon as one thinks about PD7.1, however, one realizes that it is insufficient. Most intentional verbs [17] can take expressions as objects which denote not propositions, but objects. Thus we have not only "John remembers that Mary is sick and Paul remembers it too", but also "John remembers Bill's boat and Paul remembers it too". An example of such a use of "know" is the following:

A: ... And they started driving and I don't know if you know San Gabriel Valley where Crystal Lake is?

B: I don't know the area too well, my dear.

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

However, if one restricts PD7 to pronominal expressions, then it can be defended on the basis of a frequency argument. If one examines the occurences of the verbs "know" and "believe" followed by "it" or "that" in the dialogues which are presently on line, one finds that the great majority are instances of PREF phenomena. A case by case study of intentional verbs is required here. But for the moment, let us amend and restrict PD7:

(PD7.2)

If a pronominal expression is the object of one of the two intentional verbs "know" and "believe", then it is a PREXPR.

One last clue that we can propose for the detection of pronominal PREXPRs is that the pronoun "so", when it is the object of a verb, is always a PREXPR (as far as I have been able to determine). Examples are: "Is John sick? - I think so." and "I hope that Mary passed her exam. - I hope so too."

(PD8.1)

All occurences of the pronoun "so" as the objects of

[17] On intentional verbs and their logical peculiarities, see the appendix on intentionality.

verbs other than the pro-verb "did" are PREXPRs.
This clue is infrequent but very reliable.

Let us now turn to the consideration of some examples of PREF which involve pronouns. Example (9):

A: Well, *you might have saved your life* (1), my dear Caty.

B: I know *that* (2).

A: Yeh, *you might all have been wiped out in a drunken car accident if you hadn't done that.* (3)

B: I know *it* (4).

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

Both (2) and (4) are marked as PREXPRs by rule PD7.2. The problem of candidate detection and selection now arises. We perceive (1) to be co-referential with (2) and (3) to be co-referential with (4). This suggests the following pair of blatantly rudimentary rules of candidate detection (PCD) and candidate selection (PCS):

(PCD1.1)
The candidates for co-referentiality with a pronominal PREXPR is the set of all sentences and nominalised sentences at a distance of m sentences from the PREXPR in question (before or after), where m is some search parameter.

(PCS1.1)
Select the first preceding sentence or nominalized sentence as being co-referential with a pronominal PREXPR.

These two rules seem to work in our next example (10):

A: Very briefly, I've had a lot of pain for six weeks and diagnosed more or less as a dislocated disc. Now, what's your feelings? In the first place, they said cancer - maybe. So, I had a lot of tests done, but now, since I changed doctors, he says there is, well, *there's this blood test called C.E.A., which will tell if there is cancer anywhere in your system* (1). And I could hardly believe *it* (2). Now, is there such a test?

(REAL - ORAL - SOURCE: MEDICAL-CENTERED.PROTOCOL)

And again in the following examples, drawn from the same source and which we will bring together as example (11):

A: My wife knows I wear my cap the way I like to. And I know what's good for my wife, as well as for everybody else. *I keep my business to myself, without any need of those who wear feathers in their caps.* (1) And everybody in these parts knows *it* (2), thank the Lord!

A: *Her husband arrived only this morning.* (1)

B: Oh, oh, you know *that* (2) too? Bravo!

A: *You're out of your mind!* (1)

B: Yes, *it's* (2) true! I'm out of my mind!

A: I'm going home because my husband's on my mind. *I didn't see him in church.* (1)

B: Don't think of *that* (2). He'll be along to the square.

(LITERARY - WRITTEN - SOURCE: CAVALLERIA RUSTICANA)

These rules will also operate satisfactorily on the following example (11) if suoplemented by the general principles for pronoun resolution GR1 and GR2 set forth above in section II.1:

A: I heard them say *that she had been a loose woman* (1)! I don't want to believe *it* (2) - I still don't believe *it* (3) - but I can't help feeling that *it* (4) is true. Everything points to *it* (5) - and I feel ashamed, mortified! Ashamed to show myself in her company. Everybody seems to be staring at us - I seem to feel the men ogling us! *It's* (6) frightful! But can *it* (7) really be true? Do you think *it* (8) can be true? Tell me!

(LITERARY - WRITTEN - SOURCE: MOTHERLOVE)

If the search parameter m of the general rules is sufficiently large, the expressions (2)-(8) would be determined to be co-referential with (1). Some disatisfaction might be felt with this result in regard to (6), however, since what is said to be frightful is perhaps the speaker's shame and embarassment rather than (or perhaps as well as) the purported fact that she (the mother) had been a loose woman. Intuitions are not very clear on this point, and the question can be answered either way with little impact on the dialogue analysis in

this particular case. The problem, though not very serious, signals a fact that we will return to below, viz. the fact that pronominal PREXPRs which have attitudinal adjectives predicated of them are more difficult to select candidates for than pronouns which have, say, logical predicates.

For the moment, let us comment furthur on PCS1.1 It is unclear just what we mean by "sentence" in that rule: do we mean literally a string of words ending with a period, or a sentential constituent, of which there may be several in a sentence in the literal sense? The latter interpretation seems required by examples such as the following (13):

A: Sorry to bother you, but *someone seems to have changed one of our passwords* (1) and no one here knows anything about *it* (2).

(REAL - WRITTEN - SOURCE: OC133.PROTOCOL)

and the following example (14):

A: Fred, *in about 4 minutes, we're going to hand you over to a different communications site, and it's going to take us about a minute or so to re-establish uplink* (1), so you can be prepared for *that* (2).

(REAL - ORAL - SOURCE: APOLLO-13/PAGE 379.PROTOCOL)

Secondly, we must ask how PCS1.1 fares when the preceding sentence is in the interrogative mood. There are some cases involving the PREXPR "so" in which that pronoun is involved in a PREF with the declarative transform of the preceding interrogative sentence: "Is Mary sick? - I believe/think so." Such cases only occur when the preceding question is not of the WH-type: consider the absurdity of "Who is the President? - I believe so." Apart from the special "so" cases, there seem to be two other kinds of cases, illustrated by (a) and (b):

(a) Is Mary sick?
I don't know that.

How many feet are there in a meter?
Oh, I learned that in school.

(b) Is Mary sick?
What makes you ask that?

How many feet are there in a meter?
I often wonder about that myself.

In cases of type (a), the pronoun stands in for an answer to the preceding question, whereas in cases of type (b), the pronoun stands in for the preceding question itself. It is difficult to find a principle which would distinguish between the two cases. One possible solution would focus upon the role which the verbs in the verb-phrases containing the pronouns usually play with regard to presuppositions. This is a problem I hope to do more work on; at present I can only pose it.

Thirdly, it is clear that PCS1 is biased in favor of backward pronominalization. It will give false results in cases such as the following example (16):

A: Just what are you trying to tell me, young lady?

B: What I want to tell you is *this* (1) : *Your daughter has an opportunity to come out among people - and, perhaps, either to advance her career and gain recognition, or to become engaged to and marry a young man of a good, respectable family...* (2)

(LITERARY - WRITTEN - SOURCE: MOTHERLOVE)

PCS1 should therefore be modified so as to allow for forward propositional pronominalization, at least by treating "this", as opposed to "it" or "that", as a cue for such a forward direction.

Lastly, there are clearly cases in which the rule according to which one should select the preceding sentential clause or nominalized sentential clause would fail, such as example (17) below:

A: *The other one is an L.V.N. down at Permanente on Sunset Blvd.* (1)

B: L.V.N. ... That's lanky vertiginous nurse?

A: Right.

B: I think *that* (2) 's terrific. Listen, what are you getting Sam for your 29th anniversary?

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

Such a case might be handled by modifying PCS1 so that those candidates which were involved in "correction-actions" (in a large sense) would not be selected. Such a rule would be difficultly operationalizable, however, and given the frequency with which the present focus on the preceding sentential clause or nominalized sentential clause proves itself to be useful, it should probably be retained.

Having sketched above the very beginning of procedures for detecting PREF in relatively tractable cases, I would like to list some of the difficulties found in more unmanageable cases.

The first major problem which I see turns on the fact that we can say many things about events, states of affairs, reported speech and intentional objects. Some comments, such as comments about whether or not we believe that they obtain, or about our judgments of the truth or falsity of propositions which convey information about them, clearly apply only to propositions. But many other comments could equally well be made about physical objects. And when such comments are made, they do not provide us with

any clear way of determining that we are faced with a PREXPR. This is especially true of attitudinal comments. Consider example (18):

A: One is out in the garage. *She's a teacher, and she's got all her things out there and she's got her little radio there.* (1)

B: Hey, *that* (2) 's neat. What's her first name?

(REAL - ORAL - SOURCE: BLIND-DATES.PROTOCOL)

The adjective "neat" can be predicated of objects and actions as well as of states of affairs. It requires much sophisticated processing to determine that (2) is probably co-referential with all of (1), rather than with, say, "her little radio". In particular, on relies on one's knowledge of the situation in which the communication is taking place, a situation in which the participants are not in the same location. Let us consider another example (19):

A: And, Aquarius, for your information *we now have 136-mile perigee.* (1) Confirmed by Doppler.

B: Okay, *136-mile perigee now.* (2) *That* (3) 's very nice.

(REAL - ORAL - SOURCE: CORRECTION-ACTIONS-CORPUS.PROTOCOL)

We intuit that (3) is a PREXPR, and that it is co-referential with (2) and therefore with (1). However, it is rather difficult to know just how we do this. It would be rash to propose a rule which made such predicates as "is nice" signals of PREF, because such predicates are probably more frequently attributed to objects than they are to states of affairs. And so one is left with routines which would mark (3) as an object reference, co-referential with the noun-phrase "136-mile perigee".

The second major difficulty one should mention is that pronominal PREXPRs often have what one might call indefinite scope. That is, they are co-referential with a large but indeterminate number of preceding or subsequent propositions. As an example of this consider the expressions (1)-(6) in the lengthy example (19) below; all are indefinite in scope in differing degrees, and several involve forward propositional pronominalization (e.g. (1)):

C
...your worst experience on a blind date - especially if your first name is ...

T
Sharon

C
Sharon, how old are you?

T
I'm 21.

C
O.K., are you blond or brunette?

T
I'm a brunette.

C
O.K. Tell me about your worst experience on a blind date, my dear.

T
Well, my worst experience happened when I was uh, well I was in college, I was going to a girl's school.

C
uh, huh ...

T
...and they used to have all these singing groups, you know, come in and entertain us.

C
Do you mean singing groups - professionals, or do you mean from fraternities?

T
No, these were professionals singing.

C
Oh, yeh. Like the Four Freshmen and people like that?

T
Yeh, well, I ...

C
Like the Four Preps.

T
I'm not going to name the group.

C
Yeh.

T
Because *IT (1)* was really strange, because there were

about 3 of us and one girl knew the group. So, you know she asked would we like to go on this date with this singing group? We thought, Oh Wow, Yes!

C
You mean, two girls go out with four guys?

T
No, there were four guys but there would be 4 of us.

C
Oh, I see.

T
So, I got 2 other friends and you know, we thought *THIS (2)* was really going to be fantastic. We were going to get drunk and have a great big dinner and really have a ball.

C
Uh!

T
And so, uh, they had a limousine to come pick us up and it took us to the motel and we started, you know, giving each other the eye and getting kind of nervous.

C
They brought you right over to the motel. They figured you were going to sign up as "groupies".

T
Exactly, I think. That's what we thought, at least.

C
And, no dinner?

T
And no dinner, yes!

C
Wow!

T
So, anyway when we got there the funniest thing happened. They sat down and they passed out these little booklets, because we went to their suite.

C
Uh huh.

T

And, they started preaching about their religion the whole three hours and we were just crawling the walls to get out.

C
It was a religious frenzy?

T
Yeh!

C
Were they sitting around in their Saffron robes?

T
Exactly!

C
And their little Chinese Temple gongs?

T
Right!

C
What a weird, what a weird, wow!

T
And, we said, well can we order a drink? And, they said, Oh, *THAT's (2)* not the way of our religion - we can't drink.

C
Oh....

T
IT (3) was really the most boring time as compared to what we were expecting. I think it would have been more fun to go thrashing the hotel room.

C
Why of course! How long did you girls hand around with these religious freaks?

T
Well, we hung around ... I guess we left around 11 and I guess we got back around 3 in the morning.

C
Oh wow, you stayed too long, honey.

T
Yeh, but we were trapped actually.

C
Yeh, well did *THAT (4)* turn you against Saffron robes permanently?

T
Oh well, no, not really but I just wasn't expecting *it* (5) at that time.

C
Yeh. *THAT (6)* a dandy. I'm delighted you called, Sharon. You really surprised me. I thought I was going to have to bleep you out.

T
No. I called you twice before, Bill.

In conclusion, we may say that PREF detection is considerably more difficult than is TREF detection. It is hoped that the few preliminary approachs presented above are useful if only because suggestive.

BIBLIOGRAPHY

AUNE, Bruce, "Statements and Propositions", Nous, Vol. 1, August 1967

BAR-HILLEL, Y., "Primary Truth-Bearers", Dialectica, Vol. 27, Nos. 3-4, 1973

BARENSE, J.G., "Identity in Indirect Discourse", Journal of Philosophy, Vol. 66, June 19 1969

BERG, Jan, "What is a Proposition?", Logique et Analyse, Vol. 10, December 1967

CARNAP, Rudolf, "Meaning and Necessity", The University of Chicago Press, Chicago, 1956

CARTWRIGHT, Richard L., "Propositions", in 'Analytical Philosophy', R.J. Buttler, ed., Barnes Noble, New York, 1962

CARTWRIGHT, R., "Propositions Again", Nous, Vol. 2, August 1968

CHURCH, Alonzo, "On Carnaps's Analysis of Assertion and Belief", Analysis, Vol. 10, 1950

CHURCH, Alonzo, "The Need for Abstract Entities in Semantic Analysis", Proceedings of the American Academy of Arts and Sciences, Vol. LXXX, 1951-1954

COHEN, L. Jonathan, "Geach's Problem about Intentional Identity", Journal of Philosophy, Vol. 65, May 30, 1968

CUSHING, Steven, "The Semantics of Sentence Pronominalization", Foundations of Language, Vol. 9, No. 2, 1972

DENNETT, D.C., "Geach on Intentional Identity", Journal of Philosophy, Vol. 65, May 30, 1968

FABER, Richard N., "Statements and What is Stated", Philosophical Studies, Vol. 23, Nos. 1-2, February 1972

FITCH, Frederic B., "The Reality of Propositions", The Review of Metaphysics, Vol. 9, 1955

FREED, Bruce, "Saying of and Saying That", Journal of Philosophy, Vol. 67, No. 22, November 19, 1970

GALE, Richard M., "Propositions, Judgments, Sentences and Statements", in 'The Encyclopedia of Philosophy', Paul Edwards, ed., Macmillan The Free Press, New York, 1967

BIBLIOGRAPHY

GARNER, Richard T., "Lemmon on Sentences, Statements and Propositions", Analysis, Vol. 30, No. 3, January 1970

GEACH, P.T., "Intentional Identity", Journal of Philosophy, Vol. 64, October 26, 1967

GEACH, Peter, "Assertion", in 'Readings in the Philosophy of Language', Jay ROSENBERG and Charles TRAVIS, eds., Prentice-Hall, Englewood Cliffs, New Jersey, 1971

GENOVA, A.C., "'Jonesese' and Substitutivity", Analysis, Vol. 31, No. 3, January 1971

GINSBERG, Mitchel, "Towards Statement Identity", Critica, Vol. 4, May-September 1970

GODDARD, L., and ROUTLEY, R., "Use, Mention and Quotation", The Australasian Journal of Philosophy, Vol. 44, No. 1, May 1966

GROVER, Dorothy L., CAMP, J.L., and BELNAP, Nuel D., "A Prosentential Theory of Truth", Philosophical Studies, Vol. 27, 1975

HINTIKKA, J., "Semantics for Propositional Attitudes", in Herbert FEIGL, Wilfrid SELLARS and Keith LEHRER, 'New Readings in Philosophical Analysis', Appleton-Century-Crofts, New York, 1972

HIZ, Henry, "Referentials", Semiotica, Vol. 1, No. 2, 1969

KATZ, Jerrold J., "Semantic Theory", Harper and Row, New York, 1972

LEMMON, E.J., "Sentences, Statements and Propositions", in "British Analytical Philosophy", B.A.O. Williams Alan Montefiore, eds., Routledge Kegan Paul, London, 1966

LINCICOME, D.V.C., "Systematically Ignored Differences and the Identity of Propositions", Foundations of Language, Vol 12, No. 1, September 1974

LITHOWN, R.J. MARRAS, A., "Intentionality Without Extensionality", Philosophical Studies, Vol. 25, August 1974

LYCAN, William G., "Eternal Sentences Again", Philosophical Studies, Vol. 26, Nos. 5-6, December 1974

MARRAS, A., "Intentionality and Cognitive Sentences", Philosophy and Phenomenological Research, Vol. 29, No. 2, December 1968

MORICK, Harold, "Intentionality, Intensionality and The Psychological", Analysis, Vol. 32, No. 2, December 1971

NOLAN, Rita, "Truth and Sentences", Mind, Vol. 78, No. 312, October 1969

O'CONNOR, D.J., "Tests for Intentionality", American Philosophical Quarterly, Vol. 4, April 1967

PAGE, Edgar, "Reference and Propositional Identity", Philosophical Review, Vol. 79, January 1970 PAP, Arthur, "Propositions, Sentences and the Semantic Definition of Truth", Theoria, Vol. 20, 1954

PARTEE, Barbara Hall, "The Syntax and Semantics of Quotation", in 'A Festschrift for Morris Halle', Stephen R. ANDERSON and Paul KIPARSKY, eds., Holt, Rinehart and Winston, New York, 1973

PITCHER, George, "Propositions and the Correspondence Theory of Truth", in 'Readings in the Philosophy of Language', Jay ROSENBERG and Charles TRAVIS, eds., Prentice-Hall, Englewood Cliffs, New Jersey, 1971

POSPESEL, Howard, "The Non-existence of Propositions", Monist, Vol. 53, No. 2, April 1969

REINHARDT, L.R., "Propositions and Speech Acts", Mind, Vol. 76, April 1967

RICHARDS, Barry, "A Point of Reference", Synthese, Vol. 28, 1974

RYLE, Gilbert, "Are There Propositions?", Proceedings of the Aristotelian Society, Vol. 30, 1929-1930

SAYWARD, Charles, "Propositions and Eternal Sentences", Mind, Vol. 77, October 1968

SCHEFFLER, Israel, "An Inscriptional Approach to Indirect Quotation", Analysis, Vol. 14, 1954

SCHEFFLER, Israel, "Inscriptionalism and Indirect Quotation", Analysis, Vol. 19, 1958

SCHEFFLER, Israel, "Anatomy of Inquiry", Alfred A. Knopf, New York, 1963

SNYDER, Aaron, "On What Has Truth Value", Philosophical

SOSA, Ernest, "Propositional Attitudes de Dicto and de Re", Journal of Philosophy, Vol. 67, No. 21, November 5, 1970

STALNAKER, Robert C., "Wallace on Propositional Attitudes", Journal of Philosophy, Vol. 66, No. 22, November 20, 1969

THOMSON, James F., "Truth-Bearers and the Trouble about Propositions", Journal of Philosophy, Vol. 66, No. 21, November 6, 1969

TOMBERLIN, James, "Statements and Their Idenity Conditions", Logique et Analyse, Vol. 11, 1968

BIBLIOGRAPHY

WALLACE, John, "Propositional Attitudes and Identity", Journal of Philosophy, Vol. 66, No. 6, March 27, 1969

ZEMACH, Eddy M., "Reference and Belief", Analysis, Vol. 30, No. 1, October 1969

ZIFF, Paul, "What is Said", in D.DAVIDSON and G.HARMAN, eds., "Semantics of Natural Language", D.Reidel, Dordrecht-Holland, 1972

ZWICKY, Arnold M., "On Reported Speech", in C. H. FILLMORE and D. T. LANGENDOEN, "Studies in Linguistice Semantics", Holt, Rinehart and Winston, Inc., New York, 1971

DISTRIBUTION LIST

AC-S, Education Programs
Education Center, MCDEC
Quantico, VA 22134

AFHRL-AS (Dr. G.A. Eckstrand)
Wright-Patterson AFB
Ohio 45433

ARI Field Unit - Leavenworth
P.O.Box 3122
Ft. Leavenworth, KS 66027

Advanced Research Projects Agency
Administrative Services
1400 Wilson Blvd.
Arlington, VA 22209
Attn: Ardella Holloway

[illegible] versity Library
[illegible]SE 76-443
[illegible]well AFB, AL 36112

Prof. Earl A. Alluisi
Code 287
Dept. of Psychology
Old Dominion University
Norfolk, VA 23508

Dr. Daniel Alpert
Computer-Based Education
Research Laboratory
University of Illinois
Urbana, IL 61801

Dr. John R. Anderson
Dept. of Psychology
Yale University
New Haven, CT 06520

Armed Forces Staff College
Norfolk, VA 23511
Attn: Library

Ms. Carole A. Bagley
Applications Analyst
Minnesota Educational
Computing Consortium
1925 Gather Ave.
Lauderdale, MN 55113

Dr. James Baker
U.S. Army Research Institute
1300 Wilson Blvd.
Arlington, VA 22209

Dr. M.A. Bertin, Scientific Director
Office of Naval Research
Scientific Liaison Group-Tokyo
American Embassy
APO San Francisco 96503

Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

Dr. John Brackett
SofTech
460 Totten Pond Rd.
Waltham, MA 02154

Dr. Robert K. Branson
1A Tully Bldg.
Florida State University
Tallahassee, FL 32306

Dr. John Seeley Brown
Bolt, Beranek, and Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Dr. Victor Bunderson
Institute for Computer Uses
in Education
355 EDLC Brigham Young University
Provo, UT 84601

Dr. Ronald P. Carver
School of Education
University of Missouri-Kansas City
5100 Rockhill Rd.
Kansas City, MO 64110

Jacklyn Caselli
ERIC Clearinghouse on
Information Resources
Stanford University
School of Education - SCRUT
Stanford, CA 94305

Century Research Corporation
4113 Lee Highway
Arlington, VA 22207

Chairman, Leadership & Law Dept.
Div. of Professional Development
U.S. Naval Academy
Annapolis, MD 21402

Chief of Naval Education and
Training Support (DIAI
Pensacola, FL 32509

Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

Dr. Allan M. Collins
Bolt, Beranek and Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Dr. John J. Collins
Essex Corporation
6305 Caminito Estrellado
San Diego, CA 92120

Commandant
U.S. Army Institute of Administration
Attn: EA
Fort Benjamin Harrison, IN 46216

Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Commanding Officer
Naval Health Research Center
San Diego, CA 92152
Attn: Library

Capt. H.J. Connery, USN
Navy Medical R&D Command
NNMC, Bethesda, MD 20014

Dr. T.E. Cotterman
AFHRL-ASR
Wright Patterson AFB
Ohio 45433

Mr. Joseph J. Cowan, C[illegible]
Psychological Research Br[illegible]
(G-P-1-62)
U.S. Coast Guard Headq[illegible]rs
Washington, DC 20590

DCDR, USAADHINCEN
Bldg. #1, A310
Attn: AT21-OED Library
Ft. Benjamin Harrison, IN 46216

Dr. Ruth Day
Dept. of Psychology
Yale University
2 Hillhouse Ave.
New Haven, CT 06520

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

Director, Office of Manpower
Utilization
HQ, Marine Corps (Code MPU)
BCB, Building 2009
Quantico, VA 22134

Director, Management Information
Systems Office
OSD, M&RA
Room 3B917, The Pentagon
Washington, DC 20301

Dr. Donald Dansereau
Dept. of Psychology
Texas Christian University
Fort Worth, TX 76129

Dr. Ralph Dusek
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

ERIC Facility-Acquisitions
4833 Rugby Ave.
Bethesda, MD 20014

Dr. John Eschenbrenner
McDonnell Douglas Astronautics
Company-East
PO Box 30204
St. Louis, MO 80230

Major I. N. Evonic
Canadian Forces Personnel
Aplied Research Unit
1107 Avenue Road
Toronto, Ontario, CANADA

Dr. Beatrice Farr
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Marshall J. Farr, Director (41
Personnel & Training Research Programs
Office of Navy Research (Code 458)
Arlington, VA 22217

Dr. James M. Ferstl
Employee Development:
Training Technologist
Bureau of Training
U.S. Civil Service Commission
Washington, DC 20415

Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Larry Francis
University of Illinois
Computer Based Educatonal
Research Laboratory
Champaign, IL 61801

Dr. John R. Frederikson
Bolt, Beranek & Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Dr. Alfred R. Fregly
AFOSR-NL, Building 410
Bolling AFB, DC 20332

Dr. Frederick C. Frick
MIT Lincoln Laboratory
Room D 268
PO Box 73
Lexington, MA 02173

Dr. Vernon S. Gerlach
College of Education
146 Payne Bldg. B
Arizona State University
Tempe, AZ 85281

Dr. Robert Glaser
Co-Director
University of Pittsburgh
3939 O'Hara St.
Pittsburgh, PA 15213

Capt. D.M. Gragg, MC, USN
Head, Section on Medical Education
Uniformed Services Univ. of
the Health Sciences
6917 Arlington Road
Bethesda, MD 20014

Dr. Duncan Hanser
School of Education
Memphis State University
Memphis, TN 38118

Dr. Frank J. Harris
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Col. Frank Hart, Director
Training Management Institute
U.S. Army, Bldg. 1725
Fort Eustis, VA 23604

Col. G.B. Howard
U.S. Army
Training Support Activity
Fort Eustis, VA 23604

HumRRO-Columbus Office
Suite 23
2601 Cross Country Dr.
Clumbus, GA 31906

HumRRO-Ft. Knox Office
P.O. Box 293
Fort Knox, KY 40121

HumRRO-Western Division
27857 Berwick Drive
Carmel, CA 93921
ATTN: Library

Human Resources Research Organizatio
400 Plaza Bldg.
Pace Blvd. at Fairfield Drive
Pensacola, FL 32505

Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105.

Instructional Technology Branch
AFHRL
Lowry AFB, CO 80230

Dr. Edgar Johnson
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Wilson A. Judd
McDonnell-Douglas Astronautics Co.
East
Lowry AFB
Denver, CO 80230

Dr. Arnold F. Kanarick
Honeywell, Inc.
2600 Ridgeway Pkwy.
Minneapolis, MN 55413

Dr. Milton S. Katz, Chief
Individual Training & Performance
Evaluation Technical Area
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Steven W. Keele
Dept. of Psychology
University of Oregon
Eugene, OR 97403

Dr. David Klahr
Dept. of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. Alma E. Lantz
University of Denver
Denver Research Institute
Industrial Economics Division
Denver, CO 80210

Dr. Sylvia R. Mayer (MCIT)
HQ Electronic System Division
LG Hansom Field
Bedford, MA 01730

William J. McLaurin
Room 301
Internal Revenue Service
2221 Jefferson Davis Hwy.
Arlington, VA 22202

Dr. Donald E. Meyer
U.S. Air Force
ATC-XPTD
Randolph AFB, TX 78148

Military Assistant for Human Resources
Office of the Director of Defense
Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

Dr. Andrew R. Molnar
Science Education Dev. & Res.
National Science Foundation
Washington, DC 20550

Dr. Ross L. Morgan (AFHRL-ASR)
Wright-Patterson AFB
Ohio 45433

Dr. Lee Munday
Houghton Mifflin Co.
P.O. Box 1970
Iowa City, IA 52240

Navy Personnel R&D Center
Code 306
San Diego, CA 92152
Attn: Dr. James McGrath

Navy Personnel R&D Center
San Diego, CA 92152
Attn: Dr. J.D. Fletcher

Navy Personnel R&D Center
Code 1
San Diego, CA 92152
Attn: Dr. James McGrath

Navy Personnel R&D Center
San Diego, CA 92152
Attn: Library

Dr. Leon Nawrocki
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Donald A. Norman
Dept. of Psychology C-009
University of California, San Diego
La Jolla, CA 92093

Dr. Harold F. O'Neil, Jr.
Advanced Research Projects Agency
Cybernetics Technology, Room 623
1400 Wilson Blvd.
Arlington, VA 22209

Mr. Thomas C. O'Sullivan
TRAC
1220 Sunset Plaza Dr.
Los Angeles, CA 90069

ONR Branch Office
495 Summer Street
Boston, MA 2210
Attn: Dr. James Lester

ONR Branch Office
1030 East Green Street
Pasadena, CA 91101
Attn: Dr. Eugene Gloye

ONR Branch Office
536 S. Clark Street
Chicago, IL 60605
Attn: Dr. Charles E. Davis

Office of Naval Research
Code 200
Arlington, VA 22217

Mr. Luigi Petrullo
2431 N. Edgewood St.
Arlington, VA 22207

Dr. Kenneth A. Polycyn
PCR Information Sciences Co.
Communication Satellite Applications
7600 Old Springhouse Rd.
McLean, VA 22101

Principal Civilian Advisor
for Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508
Attn: Dr. William L. Maloy

R.Dir. M. Rauch
P II 4
Bundesministerium der Verteidigung
Postfach 161
53 Bonn 1, GERMANY

Research Branch
AFMPC-DPMYP
Randolph AFB, TX 78148

Dr. Joseph W. Rigney
University of So. California
Behavioral Technology Laboratories
3717 S. Grand
Los Angeles, CA 90007

Dr. Marty Rockway (AFHRL-TT)
Lowry AFB
Colorado 80230

Dr. Andrew M. Rose
American Institutes for Research 1055
Thomas Jefferson St. NW
Washington, DC 20007

Dr. Leonard L. Rosenbaum
Chairman
Dept. of Psychology
Montgomery College
Rockville, MD 20850

Dr. Worth Scanland
Chief of Naval Education & Training
NAS, Pensacola, FL 32508

Scientific Advisor to the Chief
of Naval Personnel (Pers Or)
Naval Bureau of Personnel
Room 4410, Arlington Annex
Washington, DC 20370

Dr. Robert J. Seidel
Instructional Technology Group, HumRRO
300 N. Washington St.
Alexandria, VA 22314

A.A. Sjoholm, Head Technical Support
Navy Personnel R&D Center
Code 201
San Diego, CA 92152

Dr. A. L. Slafkosky
Scientific Advisor (Code RD-1)
HQ, U.S. Marine Corps
Washington, DC 20380

Dr. Marshall S. Smith
Associate Director
NIE-OPEPA
National Institute of Education
Washington, DC 20208

Dr. Alfred F. Smode, Director
Training Analysis & Evaluation Group
Department of the Navy
Orlando, FL 32813

Dr. Richard Snow
Stanford University
School of Education
Stanford, CA 94305

LCDR J.W. Snyder, Jr.
F-14 Training Model Manager
VF-124
San Diego, CA 92025

Dr. William Strobie
McDonnell-Douglas Astronautics Co.
East
Lowry AFB
Denver, CO 80230

Dr. Persis Sturgis
Dept. of Psychology
California State University-Chico
Chico, CA 95926

Superintendent (Code 1424)
Naval Postgraduate School
Monterey, CA 93940

Technical Director
U.S. Army Research Institute for the
Behavioral & Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

Capt. Jack Thorpe, USAF
AFHRL-FTS
Williams AFB, AZ 85224

Mr. Walt W. Tornow
Control Data Corporation
Corporate Personnel Research
P.O. Box O-HQN060
Minneapolis, MN 55440

Dr. Benton J. Underwood
Dept. of Psychology
Northwestern University
Evanston, IL 60201

Dr. Carl R. Vest
Battelle Memorial Institute
Washington Operations
2030 M Street NW
Washington, DC 20036

Dr. Joseph Ward
U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

Dr. Claire E. Weinstein
Educational Psychology Dept.
University of Texas at Austin
Austin, TX 78712

Dr. David J. Weiss
Dept. of Psychology
N660 Elliott Hall
University of Minnesota
Minneapolis, MN 55455

Dr. Keith Wescourt
Dept. of Psychology
Stanford University
Stanford, CA 94305

Dr. Joseph L. Young
Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Robert Young
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

DOCUMENT RESUME

ED 136 781 IR 004 568

AUTHOR Mann, William C.; And Others
TITLE Working Papers in Dialogue Modeling, Volume 2.
INSTITUTION University of Southern California, Marina del Rey. Information Sciences Inst.
SPONS AGENCY Advanced Research Projects Agency (DOD), Washington, D.C.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO ISI/RR-77-56
PUB DATE Jan 77
CONTRACT N00014-75-C-0710; NR-154-374
NOTE 113p.; For Volume 1, see IR 004 567

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS *Artificial Intelligence; Communication (Thought Transfer); *Computational Linguistics; Computer Oriented Programs; Computers; Dialogue; *Language Research; Man Machine Systems; *Models; Programing Languages
IDENTIFIERS Dialogue Modeling

ABSTRACT

The technical working papers that comprise the two volumes of this document are related to the problem of creating a valid process model of human communication in dialogue. In Volume 2, the first paper concerns study methodology, and raises such issues as the choice between system-building and process-building, and the advantages of studying cases over studying general language use. The second is a design paper on the MATCH process of the Dialogue Model System. The MATCH program examines a node in the workspace (WS) and one in the long term memory (LTM) and decides whether the WS node can be regarded as exactly the same concept as the LTM node or as a specific instance of the general concept represented by the LTM node. The paper explores methods for making MATCH efficient and selective. The third paper concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation. The two argument forms examined are (1) "means-end" argumentation and (2) an argument from present speech-acts to statements about the present or future behavior of the author of those speech-acts.
(Author/WBC)

ARPA ORDER NO. 2930
NR 154-374

ISI/RR-77-56
January 1977

William C. Mann
Greg W. Scragg
Armar A. Archbold

Working Papers in Dialogue Modeling, Volume 2

INFORMATION SCIENCES INSTITUTE
4676 Admiralty Way/Marina del Rey/California 90291
(213) 822-1511
UNIVERSITY OF SOUTHERN CALIFORNIA

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ISI/RR-77-56	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) WORKING PAPERS IN DIALOGUE MODELING - Volume 2		5. TYPE OF REPORT & PERIOD COVERED Research 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William C. Mann, Greg Scragg, Armar A. Archbold		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0710
9. PERFORMING ORGANIZATION NAME AND ADDRESS USC/Information Sciences Institute 4676 Admiralty Way Marina del Rey, California 90291		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N . RR042-06-01 RR042-06 NR154-374
11. CONTROLLING OFFICE NAME AND ADDRESS Cybernetics Technology Office Advanced Research Projects Agency 1400 Wilson Blvd., Arlington, VA 22209		12. REPORT DATE January 1977 13. NUMBER OF PAGES 108
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Personnel and Training Research Programs Office of Naval Research - Code 458 800 No. Quincy St. Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

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)

artificial intelligence, cognitive psychology, computer dialogue, linguistic comprehension, process model, natural language processing, persuasion, argument, research methodology.

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

(OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

The five technical working papers that compose this document (which appears in two volumes) were prepared as part of the Dialogue Modeling Proje-t at ISI. Though diverse in scope, all are related to the problem of creating a valid process model of human communication in dialogue. All are unpublished and all but one are in a form intended for internal use by the project team; however, they are of interest beyond the boundaries of the project and have implications for related work in modeling human communication.

In Volume I both papers are on reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The other paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena and for detecting instances of them.

In Volume 2 the first paper concerns study methodology. It raises some of the following issues: how to choose between system-building and process-building, why studying cases is preferable to implementing general language-use functions as programs, how to control ad-hocness of results, why it is important to orient toward communication phenomena (in contrast to form phenomena) when studying natural language. The second is a design paper on the Match process of the Dialogue Model System, exploring methods for making it efficient and selective in its actions. The third concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation.

Each working paper appears with its original abstract or introduction.

ARPA ORDER NO. 2930
NR 154-374

ISI/RR-77-56
January 1977

William C. Mann
Greg W. Scragg
Armar A. Archbold

Working Papers in Dialogue Modeling, Volume 2

INFORMATION SCIENCES INSTITUTE
4676 Admiralty Way/Marina del Rey/California 90291
(213) 822-1511
UNIVERSITY OF SOUTHERN CALIFORNIA

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.

TABLE OF CONTENTS

ABSTRACT

The five technical working papers that compose this document (which appears in two volumes) were prepared as part of the Dialogue Modeling Project at ISI. Though diverse in scope, all are related to the problem of creating a valid process model of human communication in dialogue. All are unpublished and all but one are in a form intended for internal use by the project team; however, they are of interest beyond the boundaries of the project and have implications for related work in modeling human communication.

In Volume 1 both papers are on reference as a phenomenon in text. The first surveys reference identification and resolution methods in various existing natural language processors. The other paper explores the broader problem of reference, focusing on text reference and propositional reference. It develops problems and proposals for defining these categories of reference phenomena and for detecting instances of them.

In Volume 2 the first paper concerns study methodology. It raises some of the following issues: how to choose between system-building and process-building, why studying cases is preferable to implementing general language-use functions as programs, how to control ad-hocness of results, why it is important to orient toward communication phenomena (in contrast to form phenomena) when studying natural language. The second is a design paper on the Match process of the Dialogue Model System, exploring methods for making it efficient and selective in its actions. The third concerns the structure of persuasion dialogues, in particular how justification of actions appears in argumentation.

Each working paper appears with its original abstract or introduction.

Working Papers in Dialogue Modeling

Volume 2 - Section 1

Improving Methodology in Natural Language Processing

William C. Mann

Original Date: November 1975

Improving Methodology in Natural Language Processing

William C. Mann
USC Information Sciences Institute
Marina Del Rey, California

SCOPE

This is a position paper on understanding and improving the current styles and methods of scientific work in the application of computers to texts composed of elements from human languages, such as stories, dialogues and sentences. It deals only with kinds of research in which acoustic issues are secondary or absent. It is written specifically to precede discussion at the Workshop on Technical Issues in Natural Language Processing.

There are various orientations toward value that tend to get assumed rather than discussed at this point. They need not conflict, but some selectivity is necessary. Very roughly, there is an orientation toward understanding and scientific knowledge, and there is an orientation toward application and practical use. Many people regard understanding as a nearly-necessary prerequisite to practical accomplishment. That's the view in this paper, so we therefore concentrate on scientific values without denying the others.

There is a great diversity of activities that are carried out by recognizable methods, for which serious questions of methodology could be raised. There are tool-building and laboratory setup activities. We do not build linear accelerators or observatories, but we put large efforts into tools anyway. There are speculative and exploratory activities that influence the course of later, more formal work. Choice of phenomena to study is an absolutely crucial one of these activities. There are administrative activities for which methods are important. Staffing and seeking funds are also vital. All of these anticipate and support the creation of specific results and are vital to success.

The activities that produce the knowledge that keeps the work going are of a different kind. IT IS THESE CONSUMMATORY ACTIVITIES THAT I FOCUS ON HERE, TO THE EXCLUSION OF ALL THE OTHERS.

CONSEQUENCES OF METHODOLOGY CHOICE

We are currently at a crucial stage in the development of methodology, since we have a significant history of experience, but a great deal of remaining flexibility. For better or for worse, the methodological choices made in the next few years by our present leaders are likely to be with us for a very long time. The formal result-producing style that we adopt is particularly crucial for two reasons - first, because it ends up being the least flexible set of precedents, perhaps with the exception of basic presuppositions, and second, because it produces a strong final filtering effect on the results. The

adoption of a statistical hypothesis evaluation framework leads to different kinds of results. Likewise, our formal approach will produce its own kind of results and inherent limitations. So, we must pay careful attention to our current style.

My general attitude is that current methods can be very significantly improved, and that doing so will have a very high payoff with benefits far beyond the improvements to present and contemplated efforts. The methods currently in use are under-examined and poorly understood, and traditions are still weak enough to allow changes. There are attractive alternatives to many common practices.

PRESENT ADVANTAGES

Of the great diversity of approaches to language, the process approach represented at the workshop is uniquely capable. The two key methodological problems in the study of language over the last 2,500 years or so have been the problem of rigor and the problem of complexity. The problem of rigor in the use of natural language led to formal logics and to Godel. The problem of complexity has led to various strong reductions on the general phenomena, with tools such as the Osgood Semantic Differential, or paired-associate tests. Sequential-order phenomena and individual use of language tend to get badly obscured.

Process theory approaches the problem of rigor with methods by which process specifications are made very explicit. It approaches the problem of complexity with computers, that can hold and make use of very large numbers of processes at once. The compatibility and effective coverage of large collections of hypotheses can now actually be tested.

These are exciting, reorienting advantages that make me prefer the process approach to any other, to hold high hopes for its success, and to want it to be built on good foundations.

WHAT MAKES A DIFFERENCE?

What do we want out of our methodology? Three characteristics of a methodology are particularly important :

reliability
efficiency
integrative power

Reliability encompasses all of those things that make experiments trustworthy at face value, including repeatability, clarity of definition and freedom from various kinds of circumstantial effects that might be responsible for success. Efficiency addresses the effort required to achieve particular results. (You don't plan to do basic genetics studies on elephants; you may prefer fruit-flies as subjects.) It deals not only with the costs of performing the work, but with support costs as well. Integrative power involves the scope of the theories, what diversity of phenomena they cover, what subtheories they coordinate, what kinds of investigations they facilitate.

In order to discuss current practices we need some representative example. The one here is deliberately simple and not identified with a particular development effort. However it is composed of elements that seem to be widely used.

EXAMPLE OF A NATURAL LANGUAGE PROJECT

Step 1: Select a phenomenon: CONTRADICTION

Step 2: Select an input form: ENGLISH SENTENCES

Step 3: Select an output form: ENGLISH SENTENCES THAT CONTRADICT THE INPUT SENTENCES

Step 4: Design and draft a program in the local language: MEGALISP

Step 5: Debug on examples of opportunity, selected to exercise the code.

Step 6: Publish: "CONTRADICTION IN NATURAL LANGUAGE" by Leader and Worker.

SOME STRENGTHS IN CURRENT PRACTICE

We should hold on to the distinctive strengths of our methods in any changes we plan. These strengths are generally direct classic consequences of the use of computers to hold models:

Complexity of data and theory is easy to accommodate.

Time sequences and dependencies are preserved.

A diversity of hypotheses can be applied and tested for consistency in each experiment.

All of these have to do with integrative power, and on this dimension we are, at least potentially, in very good shape.

SOME WEAKNESSES

We have some serious problems. Here are some recurring problems with the FORM of the work:

1. Single experiments often take years to execute.

2. The activity is often treated as programming and program documentation rather than science. The consequences are generally that the

data are poorly identified and poorly chosen, the status of the programs as theory is not clear, the business of making clear theoretical claims is neglected, and the relevance of the activity to existing theories that are not programs is never established. The remainder of science is thus cut off, and left wondering whether we are into science at all.

3. The attempt to perform a general transaction, such as Sentence:Contradiction, strongly limits the complexity of the input that gets actually addressed, with the result that significant phenomena are missed. The effects of prior context, speakers' goals, tacit mutual knowledge of speaker and hearer are often attenuated by the attempt to be general.

4. The unit of production is a system. Whole systems are difficult to disseminate and difficult to judge as scientific hypotheses, and are not generally understood or appreciated by non-programming scientists.

5. Coping with ad-hocness is a problem: The system runs the examples, but what else it will do is unclear, or, the degree of tuning to the examples is unclear, or, the representativeness of the examples is unclear, or, the rightness of the answers is only established intuitively.

We have problems with the CONTENT of the work. There are many problems, which may be a healthy condition, but I want to attend to just one that seems to be otherwise.

In the common notion, a natural language is a scheme of communication that people use. The fact that a language is used to communicate has strong consequences. For example, as languages change, their adequacy for communication must be maintained.

The communication properties of language are being ignored in a wide variety of approaches, including processing approaches. Often, it is outside of the paradigmatic scope of the studies.

Communication deals with changing correspondences between the knowledge of one individual or system and the knowledge of another. It is more than relations between strings and strings, or relations between strings and generators of strings (syntax). It is more than relations between strings and a world or a data base (semantics). Communication involves two active processors, and an adequate theory of language will specify some consequences of that fact. By restricting the view to a single processor (or less), I suspect that we are cutting ourselves off from the organizing principles that produce the regularities that we are trying to study.

Some of the changes of style that I would suggest are implicit in the identifications of the problems cited above:

Design clear data collection methods.

State theoretical claims that are distinct from the programs. (The claims may still contain algorithms, of course.)

Decommit from attempts to be general, except where an empirical demonstration of generality is included in the work.

Shift from focus on systems to focus on algorithms.

Do something to drastically shorten the period required to do single experiments.

Beyond these suggestions, the special advantages of case analysis should be considered.

CASE ANALYSIS AS THE BASIS FOR AN ALTERNATE PROCESSING METHODOLOGY

Case analysis as a basic scientific activity is an attractive alternative to the current methodology sketched above. How would it work?

STEPS IN A CASE-ANALYSIS-BASED DEVELOPMENT IN NATURAL LANGUAGE PROCESSING

Step 1: DATA ACQUISITION. Examples of real-world use of natural language are collected. Some are selected for detailed attention.

Step 2: PHENOMENON IDENTIFICATION: The data are annotated and scored for particular phenomena of interest. Data can be scored for several phenomena at once. Scoring is performed by people who understand the language and the circumstances of the data occurrence, and who are given explicit instructions on what to look for and how to annotate it. The result of this step is a Commentary on the data.

Examples:

a. Identify requests and judge whether they are fulfilled in running dialogue.

b. Identify repeated references to an object, action or idea in a document.

Step 3: CASE MODELING: Custom-build for this data, a new one-shot program that will take the data as input, and make entries into a simulated Hearer's Memory. The program is the Model, and its "output" is its trace.

Step 4: MODEL EVALUATION: Compare the Commentary with the execution trace of the model. For each significant event identified in the Commentary, decide whether there was a correctly corresponding event in the model's execution.

With suitable selections of phenomena for study, it is not hard to decide whether the program performed appropriately. However, a serious problem remains: a program for a single case can be entirely ad hoc. This is an advantage, in that it is certain beforehand that the program will run successfully, independent of the complexity of the phenomena. But the program may or may not have any long-term significance.

The program is composed of cooperating processes. Each process can be considered to be an over-specified hypothesis, over-specified because details such as the programming language are inessential to the corresponding functional claims about language.

> VERIFICATION STEP: In order to meet the ad-hocness problem, these hypotheses must be verified by repeated application to a diversity of cases. The experiment steps cited above must be repeated, and their results compared. Inessential details (such as programming language and machine) may be changed, if desired, but the properties of the algorithms which form the basis for the theoretical claims of the work must be held constant.

The verified results are those algorithms that continue to work correctly, when their actions are judged against the Commentary, in model after model. These algorithms are the valuable ones both for practical application and for scientific knowledge.

ADVANTAGES OF CASE ANALYSIS METHODOLOGY

Since the data acquisition step is first rather than nearly last, stronger claims can be made for the ability to model real-world phenomena. Having the data in hand is a strong guide to implementation.

Because phenomena identification is explicit, and proceeds from explicit instructions, the resulting theory has a clear operational interpretation, since it substitutes powerful hindsight for less-powerful anticipation.

There is better control on complexity and effort, since no claims are made for the generality of the whole systems that are built. The amount of data modeled can be controlled, and a diversity of data sources can be accommodated. There is strong control over the involvement of world-knowledge in models, since most of the particulars can be anticipated by looking at the data.

The method can also be controlled by choices about whether several phenomena will be modeled in a single model or several smaller models. The smaller models are simpler, but the single model exhibits the compatibility of the parts and the consistency of the set of hypotheses.

This approach typically runs in a more data-driven, phenomena-responsive manner than a general system building approach. It avoids the situation in which system design is based on inadequate stereotypes of what might happen at the input. Programming can be more goal-directed as well, since the phenomena of interest have already been identified in the Commentary.

The problems of ad-hocness are treated explicitly, rather than being left to the suspicions of the journal readers. This facilitates representations of the degree and kinds of tests that the theories have had. (I suspect that for some current systems, many readers believe that they will only run the explanatory examples in the papers).

Finally, because of the close control and 20-20 hindsight of case analysis, more complex phenomena can be accommodated. In particular, communication between two non-identical human processors can be modeled.

AN ACTIVE EXAMPLE OF CASE MODELING METHODOLOGY

The Dialogue Process Modeling work at ISI is an active attempt to apply the ideas above with some embellishments, to real natural language processing problems. All of the recommendations are being used in identifiable ways. This work will be described in discussion at the conference as time permits.

Working Papers in Dialogue Modeling

Volume 2 - Section 2

A Selector Routine to Generate Candidates for Match

Greg W. Scragg

Original Date: September 1976

****

A SELECTOR ROUTINE TO GENERATE CANDIDATES FOR MATCH

Prepared for the Dialogue Modeling Project
at Information Sciences Institute
by
Greg W. Scragg

****

I. Introduction

Anarchistic systems such as the one being designed by the dialogue modeling group have both advantages and disadvantages. One of the major disadvantages is that because the various parts are all working independently of each other, they cannot be responsive directly to each other's needs. This necessarily causes a large amount of work, the results of which are never used. Although it is hard to get a good handle on the quantity of such 'wasted effort', it is certainly very large. The goal of the present research has been to try to get a hold of such a handle and propose ways to reduce the wasted effort while still maintaining the integrity of the anarchistic system. In particular I have focussed on the procedure MATCH. For our present purposes we will describe MATCH as a program which looks at a node in the workspace (WS) and one in the long term memory (LTM) and answers yes or no to the question: 'Can the WS node be thought of as representing either exactly the same concept as the LTM node or as a specific instance of the general concept represented by the LTM node?' (This definition is in fact somewhat more restrictive then some views of MATCH that have been expressed to me. I stick by this definition for two reasons: (i) It is more concrete and the English description seems to be more understandable and (ii) no particular instance of the more general notion has as yet been required by any part of the simulation.)

Alternatively the purpose of MATCH can be thought of as finding LTM concepts which are so similar to the WS concept that they, too, ought to be in the WS. One mind boggling problem of MATCH is that for any reasonably sized WS and LTM, the number of possible candidate pairs (#(WS)x#(LTM)) gets to be very large. If it were possible to somehow prescreen the potential pairs so as to greatly reduce the number of comparisons that must be attempted, the problems due to the quantity of wasted work would be drasticly reduced. This paper describes a procedure, SELECTOR, whose job it is to propose pairs of nodes that MATCH should attempt to find similarities between. For the present we will assume that the two procedures are written independently and the MATCH is called whenever SELECTOR proposes that two nodes should be matched.

There are two major thrusts involved with the suggestions proposed in this paper. First, criteria must be proposed for selecting the candidate nodes, which, while greatly reducing the number of MATCHs, still span (or nearly so) the set of pairs that we would like to see MATCHed (This set is very poorly defined, beyond the general description of MATCH above, but it must include at least all of the pairs MATCHed in the sample dialogue.) Second, there are methods proposed for realizing the criteria along with arguments to show that in fact the restrictions should have a significant positive effect on computation time.

A further comment or two on the purpose of having a SELECTOR function is necessary before proceeding. Within the view expressed so far, there are at least two ways of viewing the existence of SELECTOR. One is that SELECTOR exists for the sole purpose of reducing the workload of MATCH. That is SELECTOR should produce a set of pairs of nodes that have a much higher success rate than would random pairs. An alternate view of SELECTOR is that the proposed pairs might also (instead?) have a higher chance of being relevant to the system as a whole. That is SELECTOR would propose pairs that were needed by other processors in the system. Now the first of these methods would have to be based on a knowledge of the way that MATCH works. You cannot claim to be producing something of higher than average value unless you know the value system being employed. Thus SELECTOR cannot produce only successful pairs

unless it knows the criteria being used for success (in terms that it can work with such as node descriptions.) On the other hand, a SELECTOR of the second sort would need a knowledge of the overall goals of the system or at least of the current needs of the other processors. Since a basic premise of the dialogue modeling group is that the individual processors should be granted as much autonomy as possible, it seems that such an 'intelligent' SELECTOR is not what we are looking for at the present. However, from time to time, some of the proposed rules will definitely have the flavor of assumed purpose or need. Hopefully they haven't exceeded a reasonable bound.

A second requirement suggested by the above discussion is that there must necessarily be some more concrete notion of what MATCH does. Otherwise it can not be known if the suggested rules will help or not (since the only criteria for selection is whether or not a pair of nodes will succeed in MATCH and the only criteria of success of SELECTOR is if it does indeed propose a near minimal set containing all successful match pairs. Indeed it might be possible to have a SELECTOR that operated on a totally different criteria than MATCH if there was known to be a very high correspondence rate between the two sets of criteria. But since there can also be no known correspondence unless the MATCH criteria is known, this observation is of little use. For this reason, there is also a suggested MATCH routine (slightly modified from the existing routine) included in the section for procedural descriptions, along with arguments for the changes.

I have tried to separate items that I considered to be major problems with the dialogue modeling system and put them in a separate section at the end of this report. There will be times however when problems with the alignment of the various parts of the system will impact upon the SELECTOR routine. At such points it will be appropriate, indeed necessary, to discuss these problems in terms of (at least) how the parts work together.

II. BLACK BOXES

This section describes what it is that the proposed SELECTOR does as opposed to how it does it. That is, it describes which pairs of nodes will be offered as candidates to MATCH and what information is used in making these decisions. For convenience, the description is divided according to the nature of the information used to make the decisions of candidacy.

Some decisions on the eligibility of a WS (or LTM) node can be made by looking only at the WS (or LTM) and for the moment disregarding the LTM (or WS). Other decisions require that you look at the more global aspects - that is, look at the LTM and WS together. WE will first look at the isolated case. Within this case we can again divide the sorts of tests that we can make on a node into two classes. I have called the first of these classes HEADNESS and the second HISTORY.

A. HEADNESS

I call a node a headnode if and only if all static information about the node indicates that it should be a MATCH candidate. That is, headness is the property that the structure of the node corresponds to a structure that could match with something in a relevant way. The name head was chosen because it will turn out later that nodes matching the criteria for headness tend to be the heads of subtrees imbedded within the net. History on the other hand encompasses information about events and changes that have occured which might impact the likelihood of the node MATCHing. Thus the history is relevant for a given node if some event has occured which is not evident from the structure of the node alone and this event is known to be of a sort which can alter the probability of the node MATCHing, such as the fact that the node does not look the same as it did at the last invocation of MATCH.

The desired property of a headnode in the workspace is that the node correspond to a complete concept that is present in the head. This can be thought of as being analogous to complete sentences of English or well formed expressions of logic. In logic we never operate on subexpressions of asserted expressions and we won't here either. Given PvQ alone, it is not possible to deduce theorems based on Q. The nodes that fit this description are the nodes that have no incoming arcs (here - and throughout - the terms incoming and outgoing refer to the obvious ends of the relations when drawn as arcs. That is the existence of an inverse arc should not be construed as showing an incoming arc. Also, IAO, AKO and AVO are ignored, with the possible exception of AKO described in the last section.) To see that this is so, consider a node N with an incoming arc. N must be in some case relation to some predicate P represented by the node at the other end of the arc. That is, P says something about N. In particular, P might say that N is not true (P = NOT) or that N is a pattern to be watched for (P = (N=>M)) or that N is only true in certain circumstances (P = (M=>N)). For example, this eliminates such things 'the girl with no shoes' in 'John hit the girl...'. And also 'john saw Mary' in 'Bill said "John saw Mary"'. We don't really want the system to make much of such subconstructs. So conclusions about John seeing Mary should only be made within the knowledge that it was Bill who said that John hit Mary. If we want to be able to use the lower nodes for some purpose we will need either theorems for specific case (e.g. 'person said x' suggests 'x') or hope that

PROTEUS can bring the subconstruct in. For example, the fact that Bill told us X might cause X to be brought in as KNOWn if we have confidence in Bill.

This concept of headness is directly analogous to Schubert's asserted nodes. That is, a headnode represents a complete concept in the modeled head. It is not imbedded in some deeper structure which gives meaning to the node on a local basis. Schubert's intention was to be able to represent 'John hit Mary' without asserting that it is true that 'John hit Mary' as in "Bill said 'John hit Mary'". In this case he would want to assert that 'Bill said ...' Similarly here a head node is this highest level construct. This is the first criteria for SELECTOR:

S1: The nodes proposed by SELECTOR must both be headnodes.

To test out the notion of headness as a criteria for eligibility for a MATCH I checked out the 26 invocations of MATCH required for the simulation in the proposal. And I found the following observations to be relevant. MATCH was invoked 26 times in the sample dialogue. The nature of the matched items is summarized as follows:

For WS: Headnode - 23 times
Not headnode - 3 times.
For LTM: Headnode (as part of helping game) 3 times
Left hand side of rule: 17 times
Part of conjunction: 5 times
Other non headnode: 1 time.

At first glance, it appears that the headness notion has a few failings, especially in LTM. But several comments are in order, most of which will make these results seem slightly better. First, cycle 1-18.5 is in error. (It looks to me like it was entered as a last minute patch) This cycle accounted for the single unclassified LTM failure and one of the WS failures. However, the cycle was not necessary for in fact the MATCH that it performed must have been made in the process of making the MATCH in cycle 1-14 (it is a straight subpart to corresponding subpart MATCH).

The other two WS misses seem to me to fall out of a slightly inconsistent use of the KNOW predicate. In particular we have, in cycle 1-14, a MATCH of PROP, a subpart of (o know PROP), with a top level item in the helping game. The justification of having meta-predicates like KNOW was that it is important to distinguish whether PROP was known or only believed. But at this point, the importance seems to be abandoned. In general it seems to me that any construct in the workspace with "o knows ..." above it should probably be eligible. (Actually the rational of the meta-predicate KNOW seems dubious to me. See further comments in the last section). This, then, is the first proposed modification of the notion of headness: Nodes whose only incoming arc is the object link from a meta-predicate such as KNOW should also be headnodes.

Another interesting development was the frequency of matches involving one of a pair of conjuncts - 5 times matches were made between two items in WS and the two conjuncts of the left half of a theorem. More properly I suppose that is two MATCHs which in combination bring in the conjunction and then PROTEUS must bring in the theorem. To correct this the definition of headness will have to be transitive across conjunctions (i.e., if a conjunction is a headnode then so are the two conjuncts).

And of course it seems obvious that we must allow the antecedent halves of LTM rules to be headnodes. Although this takes care of the remaining cases of the simulation, it also brings up some interesting points.

1. So far DEDUCE has no rule of modus tollens. When (if) such a rule is added, it seems that not(conclusion) will also have to be a headnode. This will present some problem since this node is not necessarily present.
2. It now becomes clear that headness in the WS does not necessarily imply headness in LTM and vice verse. For example, DEDUCE will cause the consequent of a LTM theorem to become active. This will structurally be a headnode in the WS but it is not in LTM.

Finally, consider this interesting case. Suppose Fthere is a rule of the form:

(prop x)
suggests
(L said (prop x))

In this case the entire left hand side of the rule is pointed to (appears in) as an argument to the predicate 'said' of the right hand part. This suggests transitivity of headness across 'if-arg's and 'and's should be the dominant factor, outweighing the existence of arcs coming in from other predicates if such rules are ever to be permitted.

Following is an explicit summary of the rules governing headness. The rules are given in terms of WS nodes. Differences between LTM and WS Headness will be described at a later point. For the present, the definitions may be thought of as applying to either WS or LTM (but remember that these are really WS definitions and that there will be minor alterations later for the LTM.)

The examination of the simulation shows that there are additional nodes that we also want to have the property of headness. And again there is a correspondence to the English or logic usage.

Rule 0: A node is a headnode if there are no incoming arcs (other than IAO, AVO, and AKO).

Rule 1: If a headnode is a conjunction, then both of the conjuncts should be headnodes (possibly removing the original headnode from the status). This is analogous to claiming the equivalence of breaking a compound sentence into two simple sentences or applying the logical rule A&B=>A.

Rule 2:If a headnode is a meta-predicate, then the object should also be a headnode. For example, There are many nodes of the form O knows X, where in fact we want to match the node with an LTM node of the form X'.(X' being similar to X). For this reason we will want to propagate headness across knows and similar predicates (believes, perceives, etc.). Now we have to be careful here. It seems dangerous to me to have both 'O knows X' and 'X' on the list of headnodes. There are easier solutions if we get rid of the meta predicates (see section on rejected ideas)

Rule 3: Eventually we might want to similarly include the disjuncts when the head is a disjunction, the negated node if the head is a negation or the main predicate if the head is a tense (although I strongly disagree that such things as tense can be thought of as predicates). Such problems do not occur in the dialogue as of yet and we shall cross that bridge later.)

Rule 4: (LTM only) if a headnode is a rule, then the antecedent of that rule is also a headnode.

A very important problem concerning headness is that we would like some sort of guarantee that this subspace of the data base is indeed a sufficient base for spanning the entire data base. That is, how can we know that by examining only headnodes that we can find enough of the nodes that we think should match (in our intuitive notion of match) to ensure that the program as a whole won't fail at this point? I have already mentioned that these are sufficient for the sample dialogue, but there are a great many nodes in the ws and only a few of them are headnodes. Might we want to MATCH one of these non-headnodes. To get a feel that we won't, reason backwards and assume that there is a non-headnode that may be of interest. Since it is a non-headnode, there must be an incoming arc. Since there is an incoming arc, the node is a parameter for some predicate. This predicate represents the context in which the lower predicate is imbedded. The lower node will MATCH if and only if the higher node MATCHs, which means that if the analogies to natural language and logic are good, then we are safe here. (Note that, strangely, for the nodes that are headnodes by the rules above, this constraint does not seem to apply.) This inductive argument is not quite foolproof. It is conceivable that there could be a bona fide ring of incoming arcs (although I have been unable to construct one.) But barring such a ring, it should be clear that every node which is not a headnode is subordinate to a headnode and so a MATCH will be attempted between it and any node occupying a corresponding position under a headnode in the other data base (WS or LTM).

B. HISTORY

Even though a node is a headnode, it may be possible to determine that it cannot MATCH anything. In particular, if a MATCH were just unsuccessfully attempted between a headnode of WS and one of LTM (and there have been no intervening events) then it is safe to say that they won't match now. In an anarchistic system, it is very important to build some sort of recognition of this fact into the system processors so that the system does not get caught in an endless cycle of MATCHing the same pair over and over (i.e. it is important to remember your failures).

To simply remove such a node from the list of headnodes in such a circumstance is too strong a step for it is conceivable that there might occur some change in the WS that could alter the failing node. But it is possible to keep track of all relevant changes made to a node by PROTEUS, MATCH or some other processor, so that proper notice can be made and all appropriate reMATCHes can be attempted. We will see later that there is some question about the desirability of such a process. However, it doesn't seem to be of any theoretical harm in the sample dialogues and the good features probably outweigh the bad.

To help keep track of headnodes in WS that are eligible for MATCH, the system will maintain a list, WSWATCH, which should contain all WS headnodes which can be eligible. It should obey the following rules:

1. Headnodes are added to WSWATCH when they first appear in the WS.

2. Items are removed from WSWATCH when a cycle of MATCH is completed. Note that one item from WSWATCH may be attempted in MATCH with several items in LTM, removal from the list should be thought of as occuring at the completion of the attempted MATCH against all items in LTM. That is, items are removed from the list when they are found not to be MATCHable against anything.

3. If a structure is rejected because of a subnode, the structure may later become good if the troublesome node is pruned off. That is, if you think of the headnode as being connected to each subpart by a chain of relations, then if any relation on this chain is changed the path is broken. So it seems that we must save some sort of representation of the path from the headnode to the subnode. Then if a destructive (change or forget) change is made in the ws affecting the path between the headnode and the place of conflict, then the headnode should again become eligible for MATCH and be re-added to WSWATCH. All of this can be summarized as

S2: No WS node can be suggested by SELECTOR if it has been previously found to be totally unMATCHable and the subpart causing the failure has not been changed. (Note that when failure is caused by the MATCHPAIRS test, the failure point can be either of the two matched-pairs.)

There is a philosophical question about WS changes that should be discussed here. What does it mean, in common language, when nodes in WS change? Changes to a node can happen in 3 different ways: a link can be added, it can be removed or it can be replaced by a different link (although the last one can be thought of as a combination of the of the first two). If a node has previously failed in MATCH, the addition of a new link cannot enable MATCH to succeed, because the present definition of MATCH always succeeds when the only problem is lack of information. On the other hand, removal of a link can enable a previously failing MATCH to succeed if the removed link was on the chain to the failure point. Now all of this implies that the deletion of information is the relevant factor in determining MATCHing. Yet deletion is equivalent to forgetting or at least to cessation of attention. IT seems very strange that the less attention paid to a node, the more likely a MATCH is. Thus MATCH should be regarded as a sort of negative test - to succeed at MATCH means that the possibility of relevance cannot be ruled out.

A real (computational) problem also exists for the change criteria: Recall that for every MATCH failure a record must be kept of the path from the WS headnode to the failure point. Even in the restricted set of pairs that SELECTOR produces there will be a lot of MATCH failures. In addition, it is obvious that there will be a great deal of changes to WS (particularly from PROTEUS). A real worry here is that the quantity of work that it takes to check every change in WS to see if it might affect some headnode might easily exceed the advantage gained from excluding non-MATCHable headnodes. In the procedure section, a few methods for reducing this work load are given.

C. GLOBAL CORRESPONDENCES

Headness in the LTM is almost identical to that in WS with a few minor exceptions. First as we noted, antecedents of rules are heads in LTM. Second LTM nodes cannot

change with time so changes cannot alter their relevance. Third, a nice advantage can be obtained by maintaining multiple lists of headnodes. In the WS there is a single list, WSWATCH, which can be thought of as containing all WS headnodes. For LTM I am proposing that instead of a single list there should be several lists, each corresponding to a single predicate. This list can be thought of as a node appended to the description of the defining node or case description for the predicate. On the list will be each head node that is an instance of that predicate (i.e. every node in LTM with pred P attached to it where P is the predicate in question). Note that by definition, each headnode must be an instance of some predicate. So such a collection of lists can in fact span the entire LTM. We can also require that each headnode be on exactly one such list by putting on the list only nodes which are instances of the predicate itself and not of more general predicates or more specific ones either. Thus JOHN RAN HOME should appear on the list for RUN but not for TRAVEL or for SPRINT, even though in fact it is the case that 'John traveled'.

Note that this list is effectively no work to create, because every instance of a predicate should be connected to the defining node anyway and we have already determined that we could make a computationally definable notion of headness. Since all arcs have inverses the list for a predicate P, correspond roughly to the intersection of pred-c(P) and the headnode list for LTM. The rationale for dividing up the list this way comes from the global perspective. Given that we have a WS element for MATCH, we know that the candidate for the LTM element will not succeed unless certain forms in particular the pred must be of the correct class which is determinable from the WS element. It might have been possible to focus in on some other aspect of the node, say actor and require that it be similar, but PRED has the advantage that it is always present and less likely to be confused via AVO links etc. Also the possible predicates form a very specific list and the instances of a particular predicate must be of very particular forms (i.e. the lower structure is more likely to MATCH). It also makes sense to divide the LTM lists up instead of the WS list because the LTM is static but the WS is always changing.

The highest level MATCH/SELECTOR can first select a WS component from WSWATCH. The first candidates for the LTM half will be on the list corresponding to the predicate of the WS candidate, additional candidates can be found by moving up and down the AKO hierarchies. Thus the third rule for SELECTOR criteria can be given as:

S3: The elements x and y that SELECTOR proposes must have predicates that are either the same or exists in a class-superclass relationship with each other.

III. PROCEDURAL OUTLINES

There are three groups of procedures needed to implement the SELECTOR function described above. First there are the procedures that make up the SELECTOR processor itself. Second, there are a number of procedures that must be included in other parts of the system to ensure proper functioning of SELECTOR. And finally there is a description of the MATCH routine. They will be described here in the reverse of this order. For each proposed function, there will be a description of how it is to work, perhaps followed by a more precise pseudo code, which will be a cross between LISP and SOL and will assume the existence of many functions pulled from the union of these two languages or even existing elsewhere. Finally there may be a discussion about the rational for the appearance of a function or even its right to exist.

A. MATCH

MATCH really does not get a general description beyond that generally known to exist and given earlier. To reiterate, it is a function that must take a pair of nodes, one from WS and one from LTM and decide if they are in fact similar, where I take similar to mean that the work space node can be taken as refering to either the same concept as the LTM node or to an specific instance of that concept. The code is a very slightly modified version of Jim Moore's MATCH procedure.

```
0: Begin
1: If (missing x) or (missing y), then succeed;
2: If (x = ENTITY) or (y = ENTITY), then succeed;
3: Increment (match-degree);
4: If IAO(x) = y, then succeed;
5: If no x in CLASS(IAO(AVO-C(X))) = Y, then fail;
6: If corresp(X, match-pairs) = Y, then succeed;
7: If corresp(X, match-pairs) = Z ~=Y, then fail;
8: If for-all r in Intersection(R(X), R(Y))
     MATCH(r(X),r(Y)) succeeds, then MATCHPAIR(X, Y),
                                    SUCCEED;
                               else FAIL;
9: end;
```

Where the subfunctions not described below, should be thought of either as the same as in <MOORE>MATCH.PROCESSOR, or as the obvious function.

CLASS should be thought of as the union of ISA(X) [AKO*(X)], with any hypothetical instances of nodes in ISA(X). By hypothetical instance, I mean a node which seems to be in a element-set relation to the defining class, but which does not have a definable referent. In particular these are the nodes that have names like PERSON/9 and HIT/2 that exist in rules and game descriptions. To facilitate recognition, I recommend the creation of a General-element (GE) link which behaves like the AKO link except that it distinguishes the element as being hypothetical. These nodes are unrestricted images of the defining class. In a rule such as (person/4 knows x) suggests ..., person/4 should have all properties of the node person and exists as a separate node only for semantic

requirements of the net ('person knows x...' somehow seems to suggest that all people know or the general concept of PERSON knows...). (In reality I would suggest partitioned networks (see last section), but the above seems to be more in keeping with the view of how the data base will look that is currently held by most members of the project.)

CLASS(X) = conj(GE(X), CLASS(AKO(X)))

FAIL is a returning function which records on the list CHANGEWATCH, the nodes that were being attempted, and returns. Note that as the recursed MATCHs unnest, the entire chain, from failure point to headnode can be produced. The highest level MATCH can then assimilate the chain into the form needed by the history checkers described below. Note also that in the case of a failure at step 7 due to a matchpair problem, that both the node being examined an the failed matchpair will have to be added to the list. The path between each of these nodes need to be watched for changes. The one path from the node currently being MATCHed to the headnode is obtained for free as the MATCHs POP back up. But it isn't yet clear that there is any nice way to save the path to the other matchpair element that caused the failure.

The functions SUCCEED and FAIL used by the highest level MATCH must note the node x. At the end of the cycle, all flagged nodes are removed from WSWATCH. Note that the nodes are removed whether they succeeded or failed, because they were attempted and no further successes can occur. ALso note that, in this case, flagging was done because there might be several MATCH attempts for one WS node during a given cycle and this is ok. What we want to avoid is trying to MATCH in a later cycle.

R is the set of outgoing arrows only and should not include AKO, IAO or AVO. This is to help provide some assurance that the recursion of MATCH will terminate. Since recursions can now occur only on outgoing case relations, they must terminate when the end of a substructure (that indicates subordinate parts of a concept) ends. It also prevents failures in MATCH due to extraneous relations that might happen to point to a given node.

r is the node at the end of a relation R.

B. Support routines

The support routines exist primarily to continuously maintain the list WSWATCH which contains all the headnodes which are thought to be possible candidates.

1. A node should first appear on the list, at the first time that it appears in WS, if it is a headnode. To do this, we must insert into the code which creates a newnode, the subroutine which will establish a pointer from WSWATCH to the newnode. (actually at this point the node should be flagged, at the end of the cycle, it will be added to WSWATCH if it is a headnode - see more detailed description of this sort of process under 3 below.)

(or old-newnode(x)
  newlist(x))

2. SUCCEED and FAIL as described above remove nodes from the list.

3. At the point of failure of MATCH, there are four items of interest: the top level nodes in WS and LTM on which the current recursion started and the particular nodes which we are attempting to MATCH when the fail occurs. This failure is the explicit failure, not the propagated failure caused by trying to MATCH each of the subparts of the node. When a failure is discovered, these four nodes will be put on the list CHANGEWATCH to be saved. IT is probable that we need only the high level WS nodes, but for now lets save all of them since we know what they are.

4. Every action that changes the WS will check this list. In particular, FORGET and ESTABLISH or their equivalents can be rewritten to make this check. If the first of the three arguments is exactly one of the names on the list (i.e. it refers to the low level WS node that failed to match), then the high level node will be re-added to the WSWATCH. Note that it might be possible to put a qualified addition into the list so that the next cycle of MATCH would only attempt to match this newly added node against the LTM node that it was attempting to MATCH against when the original failure occured (i.e. restart the MATCH that failed).

Note that I said that only the first argument had this effect of causing a check on the list. This corresponds to the claim made elsewhere that the recursions on MATCH would only be made along outgoing arrows. Incoming arrows should be of little consequence.As noted previously there must be some way to reduce the amount of work that must be done in keeping track of the nodes that have changed since the last attempted MATCH. I mentioned above the possibility that a great many sorts of changes would make no difference. If this holds true, we can certainly make the checking routines have flags so that they can tell if the procedure causing the change is one that can make significant changes (or write separate subroutine, one set to be used in the processors that make significant changes and one set to be used in the processors that do not. But the present task is to toss up some ideas about how to reduce the amount of work required to keep track of the changes that do occur.

The simplest suggestion for getting a large reduction is to take advantage of the cyclic nature of the operation of the processors. That is, assume that the notion of cycle used in the simulation will be a valid continuing notion and that there is only one processor operating at a time and that each runs to a natural stopping point before allowing the next to run. A natural stopping point for MATCH is after an attempt to match all pairs proposed by SELECTOR. For PROTEUS it is one PROTEUS cycle. Now assume that every changing function (forget or change), instead of causing a check to see if the node is on any list of relevant node, merely marks the node as changed. At the end of the cycle, a pass can be made over all nodes comparing those that have changed with those on a single combined list of node that would make a difference. This list can be sorted, assuring that only one pass through the space of nodes is necessary. There might be an advantage to having a list of changed nodes rather than simply marks on the nodes. Then the two lists could be es. compared linearly. Also the list of nodes could, of course, have pointers from each node on the list to the headnode that it effects. This tends to also suggest that what we need is not the chains from the fail points to the headnodes, but simply each of the nodes on the path with a pointer to the headnode. This can be very easily created as the MATCH successively POPs out of its ecursion after finding a failure.

C. The SELECTOR function

The SELECTOR function itself turns out to be very simple and will work as follows:

1. Select a WS candidate from the nodes on WSWATCH (By definition of WSWATCH, these are nodes that both have the property of headness and have not been attempted in a MATCH since they were last changed or created.)

2.Select a node from the set CLASS(PRED(x)). (PRED(X) is the node pointed to by the relation PRED from the node X. Using CLASS as defined in MATCH assures that all of the candidates appearing in an AKO tree above the Y directly corresponding to X will be attempted.)

3. Repeat from 2 until exhausted.

4. Repeat from 1 until WSWATCH exhausted.

Hopefully this will be representable in LISP by not much more than:

```
(mapc WSWATCH (function (lambda x
        (mapc WATCHLIST(CLASS(PRED x)) (function (lambda y
            (MATCH x y]
```

In reality, this code should be mixed with a slightly modified version of MATCH. Together, they will form the top level MATCH. The reasons for combining are that much of the work of the first level of MATCH has already been done by SELECTOR and several of the steps of MATCH are not applicable at the highest level.

Working Papers in Dialogue Modeling

Volume 2 - Section 3

A Study of Some Argument-forms in a Persuasion-Dialogue

Armar A. Archbold

Original Date: September 1976

A STUDY OF SOME ARGUMENT-FORMS IN A PERSUASION-DIALOGUE

Armar A. Archbold
September, 1976

I. INTRODUCTION: GOALS, CONCEPTS, AND METHOD

In this paper I will examine and discuss two types of argument used with persuasive intent by participants in a dialogue (or rather, multi-logue). The multi-logue is one which took place between Henry Morgenthau Junior, then Secretary of the Treasury, and two of his aides, on the one hand, and on the other hand, two representatives of the Associated Gas and Electric Company, on May 27, 1936. The two argument-forms which I will examine are (1) "means-end" argumentation (a type of so-called practical reasoning), and (2) an argument from present speech-acts to statements about present or future behavior of the author of those speech-acts.

As described, this report involves the use of three central concepts, those of persuasion, of an argument, and of the form of an argument.

1) *Persuasion* is narrowly defined for operational reasons as *the successful or unsuccessful attempt by a speaker (the persuader) to change an interlocutor's (the persuadee's) factual or evaluative beliefs so as to increase their congruence with the beliefs of the persuader, or with beliefs the persuader does not hold but wants the persuadee to adhere to, by means of the production of arguments - an attempt which meets with some resistance on the part of the interlocutor.* Resistance on the part of the persuadee is evidenced by his producing denials or expressions of dis- belief of the thesis the persuader is attempting to induce him to adhere to, or counter-arguments designed to show that the persuader's arguments lack validity or force. This definition of persuasion is narrower than most definitions, which construe persuasion to be the attempt to change not only an interlocutor's beliefs, but also his attitudes, desires, goals, emotional states or actions. It also confines persuasive techniques to arguments, which might be loosely termed "appeals to reason", as opposed to motivational or emotional appeals.

2) An *argument* is commonly defined as a coherent series of reasons given in support of the truth, plausibility or acceptability of some claim (thesis, conclusion). An argument is a statement with the support for it, support designed to meet an actual or possible challenge to the acceptability of (legitimacy of adherence to) the statement:

"A man who makes an assertion puts forward a claim - a claim on our

attention and to our belief. Unlike one who speaks frivolously, jokingly or hypothetically (under the rubric "let us suppose"), one who plays a part or talks solely for effect..., a man who asserts something intends his statement to be taken seriously; and, if his statement is understood as an assertion, it will be so taken. ...

The claim implicit in an assertion is like a claim to a right or to a title. As with a claim to a right, though it may in the event be conceded without argument, its merits depend on the merits of the argument which could be produced in its support. Whatever the nature of the particular assertion may be... in each case we can challenge the assertion, and demand to have our attention drawn to the grounds (backing, data, facts, evidence, considerations, features) on which the merits of the assertion are to depend. We can, that is, demand an argument; and a claim need be conceded only if the argument which can be produced in its support proves to be up to standard." [Stephen Toulmin, "The Uses of Argument", Cambridge University Press, London and New York, 1958, pg. 11]

Note that (a) the permissibility of challenging assertions, or types of assertions made by types of people in types of circumstances, (b) the custom of meeting such challenges by argument, and (c) the standards by reference to which the acceptability and force of arguments are judged, all vary quite largely between cultures. [Richard D. Reike and M.O.Sillars, "Argumentation and the Decision-Making Process", John Wiley and Sons, New York, 1975, pg.2ff.]

An argument is frequently thought to be an an ordered sequence of sentence tokens or types related to one another in a manner similar to that in which the sentences (expressions) in a proof are related to each other. To prove a thesis is to show that it follows from certain previously proved or axiomatically accepted propositions by the application to these propositions of certain accepted rules of inference. The notion of a proof has been rigorously defined by logicians and mathematicians in terms of the formal, artificial languages and theories they work with. Suppose we have an artificial language L whose syntax delimits a set of well-formed formulas (or "wffs") of that language. Suppose that we furthermore have a logical theory in L, LT, which comprises (1) a set of wffs of L designated as axioms, and (2) rules of inference, which consist in a set of relations, R1...Rn, among wffs. "For each Ri there is a unique positive integer j such that, for every set of j wffs and each wff A, one can effectively decide whether the given j wffs are in the relation Ri to A, and, if so, A is called a direct consequence of the given wffs by virtue of Ri." A proof is then "a sequence of wffs, A1...An, such that, for each i, either Ai is an axiom of L or Ai is a direct consequence of the preceding wffs by virtue of one of the rules of inference." The last wff in a proof is the conclusion of a proof and is a theorem of LT. [Elliot Mendelson, "Introduction to Mathematical Logic", D.Van Nostrand Co., Princeton, New Jersey, 1964, pg. 29]

In this perspective, an argument is an ordered sequence of sentences, of the form P1...Pn,{I1...In},C, where P1...Pn are the premises of the argument (presumed to be known or adhered to at the outset), C is the conclusion of the argument, and I1...In are optional intermediate sentences which may or may not

be present. Each of the I1...In, and the sentence C, are direct consequences of some of the preceding sentences by virtue of some rule of inference applicable to natural language sentences.

The rules of inference in question are normally thought to have to be valid or at least "good" rules. A valid rule of inference is one which sanctions only truth- or acceptability-preserving inferential steps - that is, one which allows one to derive only true or acceptable conclusions from true or acceptable premises. A good though less than valid rule of inference is one which in most cases, though not always, preserves truth or acceptability.

Some who think of arguments within the paradigm of proof restrict the notion even further, so that it applies only to ordered sequences of declarative sentences, sentences which are capable of being said to be true or false. A restriction of the rules of inference involved in arguments to deductive rules often accompanies this limitation.

It is attractive to think of arguments as natural language proofs, because it seems to make the clear, powerful concepts and procedures of formal logic available for their analysis. However, such an approach is too constraining if one wants to adopt a definition of argument which would allow one to study the wide range of phenomena usually referred to by that term. Let us consider some of the restrictions mentioned above, working backwards, in inverse order of their presentation:

a. We cannot restrict arguments to derivations of conclusions which use only deductive rules of the type traditionally studied by logic, at least on the face of it, because we commonly speak of a host of non-dediuctive arguments: inductive arguments, analogical arguments, arguments from authority, practical or moral arguments, "conductive" arguments, and so on. To insist that all arguments are deductive is either to propose an unreasonably constrictive stipulatory definition or to claim the following: that either (a) all non-deductive arguments can be transformed, without distortion or loss, into deductive arguments, or (b) all non-deductive arguments should not , for a variety of reasons, be deemed arguments at all. Both statements (a) and (b) are highly questionable [cf. the excellent discussion in Carl Wellman, "Challenge and Response: Justification in Ethics", Southern Illinois University Press, Carbondale, 1971, Section One]. We will reject this "deductivism". Of course, this has the drawback that we will have to work with rules of inference that are not always truth-preserving:

> "Presumably, if a good argument has true premisses and a satisfactory inference-process it must have a true conclusion too? Unfortunately, the case is not quite so simple as this. If logicians had found their perfect theory of deductive validity and we were to agree to work within the bounds of this theory, this would, of course, be so.... But this is not the case at present, and may never be; and, in any case, there are good arguments that are not deductive. In practice, although we would want to say of a good argument that it *supports* its conclusion, it is not, as a rule, possible

> to say that it supports it beyond fear of reproach or criticism. It often occurs that there are good arguments *for* a given conclusion and also good arguments against it. We cannot demand of an argument that it be, all by itself, a knock-down one. If we did, we would risk running across a situation in which we found that there existed both a knock-down argument for a conclusion and a knock-down argument against it at the same time." [C.L.Hamblin, "Fallacies", Methuen and Co., London, 1970, pg. 232]

So we should only require that rules of inference tend to preserve truth, not that they always preserve it.

b. Within the proof-paradigm, we cannot insist that all of the expressions which make up an argument be true-or-false, declarative sentences which are used to make assertions. Many of what we would naturally call arguments contain sentences used to make value statements, requests, exhortations, commands - sentences which are either non-declarative or not used to make assertions which can be said to be true or false in the strict logical or scientific sense. Consider:

> Please take all of my bags to the station.
> This bag is one of my bags.
> (Therefore) Take this bag to the station.
>
> Shoot all traitors!
> John is a traitor.
> (Therefore) Shoot John!

For an introduction to the literature on this point, cf. [Robert P. McArthur and David Welker, "Non-Assertoric Inference", Notre Dame Journal of Formal Logic, Volume XV, Number 2, April 1974] One consequence of this point in conjunction with the preceding point (a) is that the rules of inference used in the construction of arguments are only required to tend to preserve what we have so far been calling "acceptability", which includes but is not limited to truth or plausibility. Thus if one accepts (agrees to comply with) the command which forms the first premise of the second example above ("Shoot all traitors!"), the argument shows that one should accept the command which is the conclusion of the argument ("Shoot John!"). This notion of acceptability is admittedly vague, and what it is to accept an utterance varies with the nature of the utterance - to accept an assertion is to believe it, to accept a command is to be willing to comply with it, and so on. Rather than attempt a full-scale explicitation of the concept, we will in this paper only try to render its application precise in particular specific instances, when we need to do so.

c. The notion of acceptability brings up a crucial question: just what is the nature of the components of arguments? In formal logic, these components are expression-tokens or expression-types. But the components of arguments in the everyday sense are not, properly speaking, sentences: we do not believe sentences, strictu sensu; we are not willing to comply

with imperative sentences. Rather, we understand, believe in, adhere to, etc., propositions, the meanings, readings or semantic interpretations of sentences. It follows from the remarks in (b) above that I am here using the term "proposition" in a way similar to that in which J.J.Katz uses it, so that I can say that not only declarative, but also interrogative, imperative and hortatory sentences "express" propositions. [cf. A.A.Archbold, "Text Reference and Repeated Propositional Reference: Concepts and Detection Procedures", ISI, 1975, pgs.10-16, for a brief discussion of three major approachs to the notion of a proposition.] But unlike Katz, I would like to use a notion of proposition which includes (1) information imparted by the utterance of a sentence in context and cotext about the reference of terms used (and not only about their intensions), and (2) the illocutionary force of the utterance of the sentence in context. So when I say that the components of an argument are propositions, I mean that they are units of information imparted by the utterance of sentences, which comprise both referential (extensional) , semantic (intensional) and illocutionary information. I cannot provide a rigorous account of this notion of proposition (I would be a happy philosopher indeed if I could), but I believe that it corresponds quite closely to what many workers in AI are attempting to capture in their deep- structure representations. In what follows, when I write of sentences or expressions, I will intend to refer to the propositions expressed by those sentences or expressions as utterred in context.

d. We must avoid a definition of argument which makes it necessary to say that (a) a bad argument is not an argument at all, and (b) a good argument is good in every relevant respect. [John Woods and Douglas Walton, Review of C.L.Hamblin's book "Fallacies" (op.cit.), The Journal of Critical Analysis, Volume IV, No 3, October 1972, pgs. 104-105] In other words, our definition must not incorporate our evaluative standards for good and bad arguments (it would be absurd, to mention an example of such a move, to define only the best knives to be knives). In particular:

1. We should not insist that all the rules involved in the construction of an argument be rules of inference generally held in our culture to be valid or "good". An argument may involve rules that are exceedingly unreliable, in which case it is a bad argument, but an argument nevertheless. Of course, for us to recognise an argument as such, we must perceive it to involve some rules which bear some faint resemblance to some socially practised rules, either good or bad. If we came accross the following sequence of sentences:

Employees crave recognition.
Napoleon married Josephine.
(Therefore) Supersonic flight is dangerous.

we would not regard it as an argument, but rather as some exercise in poetry or free association, because we could not imagine any rules of inference in any way resembling those we are accustomed to which would sanction such a derivation.

2. The notion of an argument as a natural language proof suggests that arguments are completely explicit: all of the premises and all of the intermediary expressions necessary for the rules which sanction the derivation of the conclusion to apply are present, as is, of course, the conclusion itself. However, many of what we naturally term arguments are not explicit in this sense; many of their premises, their intermediate expressions, and indeed sometimes their conclusions are missing (suppressed, implicit). We often judge such arguments to be "good" arguments despite these omissions, when the suppressed element are such that they can be taken for granted [C.L.Hamblin, op.cit., Chpt.7]. So if we are to adopt a definition of argument which covers both good and bad arguments, we must allow not only for missing elements which can be taken for granted, but also for missing elements which would not be ordinarily thought of as being omissible, and which we must make some effort to reconstruct. Of course, here again, if a presumed argument has gaps which we are unable to fill in despite great effort, we will decline to call it an argument at all (think of the example in (1) above).

3. The components of a proof-like argument are arranged in a definite sequential order. But a person who puts foward an argument does not always first state his premises and then his conclusion; he may state his conclusion first and then adduce premises in support of it, or he may first state some premises, then a conclusion, and then some further premises. Though a "good" argument is ordered to some extent, all that is really required of an argument is that it be a collection of statements which support or are intended by the author to support some conclusion (perhaps implicit, as we said in (2) above) by virtue of some rules of inference.

In view of the above, we might be tempted to adopt an extremely loose definition of an argument: an argument is a set of propositions from which it is possible to derive another proposition (the conclusion of the argument) by means of socially practised rules of inference which (at least) sometimes preserve acceptability and - possibly - of additional propositions not present in the set but necessary for the derivation of the conclusion. However, this definition is entirely too loose, for given the permission to bring in additional propositions, the multitude of rules available to us, and the fact that the conclusion of an argument need not, according to this definition, be specified by the argument itself, just about any arbitrary set of propositions would count as an argument.

We must constrain this definition with an eye to our research goal, which is to study arguments employed by participants in persuasion-dialogues in the attempt to change their interlocutors' beliefs. How do we recognize the presence of an argument in a dialogue? It seems to me that we start looking for arguments when we perceive that there is some proposition about which the two participants disagree, so that one participant adheres to or accepts it, and the other participant either does not adhere to it or adheres to its negation (or is perceived not to do so by the first participant). We will call such a proposition a "debate proposition", to borrow a term from the forensic literature:

> "In argumentation and debate the term *proposition* means a *statement of judgment that identifies the issues in controversy.* The advocate desires to have others accept or reject the proposition. Debate provides for organised argument for or *against* the proposition: those arguing in favor of the proposition present the affirmative side; those arguing against it, the negative side." [Austin J. Freeley, "Argumentation and Debate: Rational Decision-Making", Wadsworth Publishing Co., Belmont, California, 1976, pg.30]

The debate proposition need not be expressed explicitly; it may be implied or simply evoked by something that one of the participants has explicitly said. But it must be a subject of controversy or disagreement in the dialogue: that is, we must perceive it as being expressed, implied or evoked by one participant, and questioned or contradicted by a proposition expressed, implied or evoked by his interlocutor.

Once we have detected a debate proposition, we start looking for propositions expressed, implied or evoked by the participants which (a) could be interpreted as support for either the debate proposition or its negation, and (b) seem - given our interpretation and understanding of what is going on in the dialogue - to be intentionally adduced by the participants in support of either the debate proposition or its negation. We believe that propositions P1...Pn can support a debate proposition (or its negation) C if we can, with some reasonable amount of effort, generate an argument of the form P1...Pn,I1...Im,C , which involves some socially practised rules of inference, and where I1...Im are not outrageously implausible propositions. We believe that propositions P1...Pn are in fact intentionally adduced in support of a debate proposition (or its negation) C iff (1) they can support C, and (2) they can do so by virtue of other propositions and rules of inference which the speaker explicitly or implicitly accepts (or at the very least, which he does not explicitly or implicitly reject), or which he believes his interlocutor accepts (this is to allow for ad hominem arguments). If we find such a set of propositions, then we say that we have detected an argument in the dialogue which has been employed by one of the participants.

The question remains, however: just what kind of a thing is this argument which we have detected? On the basis of the above, we can say that *an argument put forward by a participant in a dialogue is a theoretical construct used by those who understand and analyse the dialogue, a a sequence of propositions of the form P1...Pn,I1...Im,C , where*

(a) each of the propositions I1...Im,C is a direct consequence of some set of preceding Ps and Is by virtue of some socially practised rule of inference which sometimes (at least) preserves acceptability of propositions;

(b) at least one of the propositions P1...Pn is either explicitly expressed or implied or evoked by a participant in a dialogue;

(c) all of the propositions P1...Pn,I1...Im are at least compatible with the perceived beliefs of the participant, if not expressed, implied or evoked by him, or are at least compatible with the beliefs he imputes to his interlocutor;

(d) C is a debate proposition in the dialogue, such that one participant adheres to or accepts it, whereas the other does not, or is perceived by the author of the argument not to accept it.

It might be felt that we should add that the debate proposition should be one that one of the participants wants to induce the other to adhere to. This condition is already suggested by our operational definition of persuasion, however. Persuaders seek to induce their interlocutors to change their beliefs by generating arguments in support of (debate) propositions which the persuadors accept but which their interlocutors do not (at least at the outset of persuasion-dialogues).

Though the above notion of argument as a theoretical construct is the one we will most frequently use in our analytical work, we will also need, on occasion, to refer to *the set of actual utterances in the dialogue which correspond to (express, imply or evoke) some of the elements of our theoretical construct.* This set we will refer to as the *expression of the argument*, or *expressed argument*. Each utterance in the expressed argument will be referred to as an argumentative utterance.

3) What is the *form of an argument*? It seems to me that there are two related but distinguishable notions of logical or argumentative form: the first involves classification of, and induction from, examples of naturally occurring arguments, while the second is involved in applying a formal logical theory to arguments expressed in natural language [cf. my note on logical form written for Prof. Bill Woods, April 1974]. It is the former which I would like to employ here, so I will discuss it briefly. Suppose we survey a number of arguments, and we notice that we can group these arguments into argument classes (ACs), such that all the arguments in a given class only differ from each other in some respects. For example (a tired old example), we might form an argument class which includes the following arguments:

All men are mortal.
Socrates is a man.
(Therefore) Socrates is mortal.

All children are noisy.
Kevin is a child.
(Therefore) Kevin is noisy.

Some elements appear in both of these arguments, whereas other elements vary. Suppose we represent the formal structure of these arguments by a sequence of propositions containing the constant elements and variables where the variable elements appear:

All X(plur) are A.
N is a X(sing).
N is A.

The variables stand in for expressions of certain syntactic and semantic types: X(plur) and X(sing) stand in for plural and singular nouns, A stands in for adjectives, N stands in for proper names.

Larger classes of arguments can be represented in a similar manner, using variables which stand in for a wider range of expressions. Suppose we have a class of arguments which includes the following two arguments in addition to the two already given above:

All dogs eat meat.
Fido is a dog.
(Therefore) Fido eats meat.

All atheletes have low blood pressure.
Schwarzenegger is an athlete.
(Therefore) Schwarzenegger has low blood pressure.

The formal structure of these four arguments could be represented as follows:

All X(plur) VP(plur).
N is ART X(sing).
N VP(sing).

where VP(sing) and VP(plur) stand in for verb phrases with singular and plural verbs, and ART stands in for indefinite articles.

We will say that *such representations display the form of arguments belonging to a class of arguments*, and that all of the arguments belonging to a class with a given formal representation are argument of the type of that representation.

It might be objected that this notion of argumentative form is not as useful as it should be, because it is too dependent upon surface linguistic phenomena. For instance, the last argumentative form mentioned would not include the following linguistic variant of one of the arguments mentioned in the argument class which it represents (without good reason, we feel):

Dogs all eat meat.
Fido is a dog.
(Therefore) Fido eats meat.

But this objection overlooks the fact that the expressions from which argumentative form is abstracted are not (surface) sentences, but rather expressions belonging to some language, much simpler and standardized than English, which is used by analysts to represent propositions. We form argument classes out of sets of sequences of propositions already expressed in this cannonical, deep-structure form. Thus both of the surface utterances, "All dogs eat meat" and "Dogs all eat meat" are represented by the standard form "All dogs eat meat".

Rules of inference sanction steps from some sets of propositions to others; they

have some generality because they refer to sets of propositions of the same argumentative form. They therefore use the same representation as that in which argumentative form is displayed.

In this report, we will adopt the following method when examining each of the two arguments occurring in our dialogue and which we have chosen to study. We will construct each argument progressively, formulating useful rules as we go, and briefly discussing analytical problems as they arise. More specifically, we will make the following steps for each argument:

i) distinguish the debate proposition which gave rise to, and forms the conclusion (or negation of the conclusion) of the argument; specify the utterance(s) which (a) allowed us to detect the debate proposition and/or (b) expressed, implied or evoked the debate proposition;

ii) list the argumentative utterances which we feel are made by one of the participants in order to generate the argument under study;

iii) progressively construct and display the argument itself, noting which component propositions correspond to actual argumentative utterances;

iv) make some general comments about the argument analysed, and point to various broad problems which its analysis evoked.

It should be remarked that this method is part of what might be called the synchronic (as opposed to diachronic) method of studying dialogues. The order of generation of argumentative utterances is ignored; no attempt is made to explain it. The arguments are specified after having read the dialogue from beginning to end, with the full benefit of hindsight. Knowing how arguments are evoked and even how they are expressed would be of help for predicting what might be said in certain circumstances in dialogues, but not for predicting the temporal sequence of utterances.

The entire dialogue, with lines numbered, is placed in Appendix A; sections of the dialogue relevant to the analysis will be inserted in the text when needed. It will be assumed, however, that the reader is fully familiar with the dialogue as a whole.

II. MEANS-ENDS ARGUMENTATION

In the dialogue under study, the representatives of the Associated Gas and Electric Company (henceforth abbreviated as "AGEC") and representatives of the Treasury Department ("TD") engage in several arguments. The main subject of controversy between the two sides is a recent action of the Treasury: the Treasury sent a representative to a Senate Committee to oppose the adoption of some proposed amendments to a bankruptcy law known as 77-B. The Treasury did this because adoption of these amendments would, in its view, make it impossible for the Treasury to win an on-going suit it has against AGEC. The AGEC representatives believe that this action is unintelligible (absurd, irrational, unreasonable, incomprehensible, inexplicable), and perhaps that it is also unjustified (wrong, improper, unwarrantable); the TD group holds that this action is certainly intelligible and probably also justified and proper. The Treasury side argues for their belief by showing that the action was carried out in the pursuit of a higher goal. We will, in this section, examine their argumentation, which we will call "means-ends" argumentation for the moment.

Throughout our analysis, we will treat the utterances of the representatives of AGEC, on the one hand, and of the TD, on the other, as though they were made by two persons, rather than two groups. The positions of the participants on either side seem compatible enough for this move to be justified.

i) The Debate Proposition. The debate proposition is introduced in the second turn of the dialogue by the AGEC representatives, and is re-evoked by them several times thereafter. They introduce it by means of statements which have the illocutionary force of *questions, requests for explanations and/or justifications of the action* under discussion. A list of these questions follows, along with one expression of comprehension (repetition of the question) by the Treasury side (lines 248- 250).

15 Mr. Burroughs: We have one thing on our mind and that is
16 very seriously on our mind. Last week, representative of
17 this Department appeared before a Senate Committee in
18 opposition to some legislation and the reason for the
19 opposition was primarily that the legislation, if passed,
20 would be beneficial to Associated Gas and Electric. We
21 don't understand why a Government Department, first we don't
22 understand why they appeared at all and, secondly, we don't
24 understand why they oppose the legislation because it is
25 beneficial to Associated Gas and Electric.

34
35 B: It's a very logical amendment to prevent strike suits,
36 but even if it were put in at our suggestion, I fail to see
37 why the Treasury Department should oppose legislation having
38 to do with bankruptcy cases.
39

144
145 B: I do fail to understand why any Treasury Department
146 employee should voluntarily--and I have no evidence that it
147 was not voluntarily--appear before the Senate committee and
148 oppose legislation on the ground that it would let
149 Associated Gas off.
150

248 HM: You asked why we should voluntarily appear before the
249 Committee and I answer that I am proud that our organization
250 found this thing and went up there about it.
251

279
280 B: No. I am not asking anything about the tax case. I am
281 asking why don't you want Associated Gas relieved of 77-B.
282

286
287 B: No. I am here to ask you why the Treasury Department
288 felt that it was undesirable that we should be relieved of
289 77-B proceedings and why they appeared to oppose a law which
290 would have relieved us.
291

AGEC is in effect asking the following progressively more and more specific questions:

1. Why did a representative(s) of the Treasury Department appear before a Senate Committee?

2. Why did a representative(s) of the Treasury Department appear before a Senate Committee to oppose legislation having to do with bankruptcy cases?

3. Why did a representative(s) of the Treasury Department appear before a Senate Committee to oppose legislation having to do with bankruptcy cases on the ground* that it would relieve AGEC of 77-B proceedings and be beneficial to AGEC?

The important feature of these questions is that they constitute *requests for an explanation of an act performed by a purposeful and information-processing being.* The act mentioned in 2 is more specific than the act mentioned in 1, and the act mentioned in 3 is in turn more specific than the act mentioned in 2. The act mentioned in 1 is a simple physical act (appearing); the act mentioned in 2 is an act performed with a purpose (appearing in order to oppose); the act mentioned in 3 is an act performed with a purpose and with a justificatory reason (appearing to oppose X because Y). (I am assuming here that purposes and reasons are involved in the descriptions of acts, an assumption which is debatable and undoubtedly debated in the voluminous and complex philosophical literature on the logic of our talk about actions.)

As such, these questions are not, of course, statements or claims; they might simply be utterred in order to make a request. If utterred in a certain tone of voice or in a certain context, however, they might be used to make, not just a request, but a challenge - a challenge, because they might convey in context that their speaker believes that there is no explanation, or no satisfactory explanation, for the act and is therefore explicitly questioning his interlocutor's capacity to come up with one. In the context of the present dialogue, it seems to me that these questions are in fact being used to make such a challenge.

It is highly unlikely AGEC believes at the outset that there is no explanation for the Treasury's action. Rather, as Jim Levin has pointed out, they believe that there is no good or satisfactory explanation. They probably have a rough idea of the reasons the Treasury has, and which the Treasury in fact expounds in the course of the dialogue, but they believe these reasons are inadequate. So AGEC asks these questions in order (a) to communicate their belief that the Treasury's explanation is unsatisfactory, and (b) to get the Treasury to produce an explanation which AGEC can then criticize. I believe that this argumentative, strategic use of why-questions is quite common, though probably not as widespread as their use to simply request information or explanations.*

---

*The reader may wonder why I use the phrase "one the ground that" rather than "because" or "for the reason that". In the text, we have the following phrases:

(18-20) ... the reason for the opposition was primarily that the legislation, if passed, would be beneficial to Associated Gas and Electric."

(22-25) "... we don't understand why they oppose the legislation because it is beneficial to Associated Gas and Electric."

(145-149) "I do fail to understand why any Treasury Department employee should voluntarily ... appear before the Senate Committee and oppose legislation on the ground that it would let Associated Gas off."

Lines (18-20) and (22-25) do not imply that the representative, in addition to voicing his opposition, expressed explanatory or justificatory reasons for his opposition in front of the Senate Committee. Lines (145-149), on the other hand, do imply this, at least on my interpretation in context of "on the ground that". Some may disagree. My choice of the phrase "on the ground that" is not meant to suppress this uncertainty. Its consequences will be discussed below.

---

* Many of the comments made in this and the preceeding paragraph are based on my general experience and overall impression of the dialogue, but I find myself unable to support them by specific utterances in the dialogue. In particular, I cannot tell from the dialogue precisely what AGEC's initial notion of what the Treasury's explanation was.

It is because we can interpret the use of these sentences which contain why-questions as subordinate clauses by AGEC not only as requests but additionally as challenges that we can derive debate propositions from them, viz.:

(DP1) *There is no satisfactory explanation for a representative(s) of the Treasury Department having appeared before a Senate Committee.*

(DP2) *There is no satisfactory explanation for a representative(s) of the Treasury Department having appeared before a Senate Committee to oppose legislation having to do with bankruptcy cases.*

(DP3) *There is no satisfactory explanation for a representative(s) of the Treasury Department having appeared before a Senate Committee to oppose legislation having to do with bankruptcy cases on the ground that it would relieve AGEC of 77-B proceedings and be beneficial to AGEC.*

From a logical point of view, an interesting aspect of the debate propositions (DP1)-(DP3) is that they are second-level claims: they are propositions not only about the world (an action), but also about propositions (explanations) , their existence and their relationship to the world (the action). They are claims that there are no satisfactory explanations for a certain action. It might be felt that these claims are equivalent to first-order claims that these actions are unreasonable, incomprehensible, etc.. This might be true in this particular case, given the context. However, it is certainly not true in general that "There is no (good) explanation for A" is equivalent to "A is unintelligible, incomprehensible". There may be no good explanation for the axioms of logic or science, or fundamental beliefs which we rely upon constantly in our everyday interpretations , action and experience, and yet we do not usually (unless we are philosophically inclined) find them to be unintelligible or incomprehensible - quite on the contrary, these axioms form the basis of our criteria for intelligibility or comprehensibility. We will therefore not attempt a reduction of (DP1)-(DP2) to first-order claims.

ii) The Argumentative Utterances. Below are the utterances in the dialogue by representatives of the Treasury (Morgenthau, Wideman and Oliphant) which I feel express, imply or evoke some parts of the arguments against the debate propositions. I have included some utterances by the AGEC side to provide context.

40 HM: The object is very simple. We have a suit against you
41 fellows and we certainly are not going to let a joker be put
42 into some bill which is going to make it impossible for us
43 to go through with this case.

****

51 WIDEMAN: He is attorney for the petition attorneys in the
52 77-B. Well, now, Mr. Burroughs, the Treasury has a tax

claim.

B: Yes.

W: The Treasury is interested in collecting the amount of taxes, naturally, due from Associated Gas.

B: Correct.

W: It is anxious to do that in the most expeditious way consistent with reasonably fair treatment of the Associated Gas and the stockholders of the corporation. Now there are two or three methods of collecting that tax. One is through distraint on the jeopardy assessment that has ben made and seizure of your property. The Treasury has attempted to avoid that if possible.

B: Yes.

W: A bill to foreclose the tax lien has been filed in the Collection District of New York as one more moderate method than seizure and distraint, and another probability of collecting the tax through more moderate means is through 77-B in the event they are successful.

B: I don't follow that. Why should it be through 77-B proceedings. How does that help the Treasury people?

W: It may be the most appropriate and desirable way of collecting the tax from two or three angles. One is it gives the creditors and the stockholders of Associated a look-in on the proceeding, in which the Government is collecting its tax, namely: the Government is not boffling up everything, but giving the creditors a chance to be heard, whereas if you proceed otherwise, the creditors might be left out in the cold. In 77-B the Secretary may accept less than the full amount of tax and he cannot do so under other considerations.

B: But if there is no 77-B and no trusteeship the Secretary will not have to accept less than the full amount of tax. As soon as the full amount is determined, the company will do as it always has done--pay the tax.

W: As I understand it, the Treasury has taken the position the position that is has simply because it believes that will get the same treatment in the future as it has in the past, in the matter of cooperation from the Associated, in getting information that is necessary on which to compute the tax and then collect it.

104 W: And by that you mean we have got no cooperation.
105
106 B: Is that right?
107
108 OLIPHANT: In substance.

****

151 HM: Who do you think the United States Treasury is? The
152 United States Treasury belongs to the people of the United
153 States and we are here to do our job fairly and honestly and
154 if we think that legislation, which has suddenly appeared,
155 is going to deprive the people of the United States from
156 trying a case fairly, we volunteer and go up there to see
157 that the people are protected.

****

252 W: You are not just justified, Mr. Burroughs, in saying
253 that the Government had no right to take an interest in the
254 effect of that bill on the Associated Gas case because, I
255 started to tell you, of course we can't proceed, as long as
256 77-B is going on, we can't proceed in any other way except
257 through distraint because 77-B will absorb everything.
258 Another reason why 77-B is the appropriate way to handle the
259 thing is that the Court is authorized to determine the tax,
260 if it can be done, more quickly than the Board of Tax
261 Appeals. There are many reasons why that is good machinery-
262 -the best machinery in some respects from your standpoint--
263 to determine this tax liability.
264
265 B: Isn't regular machinery set up in the Board of Tax
266 Appeals for determining liability?
267
268 W: Oh, yes.
269
270 B: Why isn't that satisfactory in our case? We have always
271 paid taxes promptly as they were determined by the Board of
272 Tax Appeals.
273
274 W: Section 77-B has the effect of preserving the assets.
275 By the time the Board of Tax Appeals gets it, there may be
276 nothing left to collect.

****

292 W: Let me give you one general answer. The stockholders
293 and all creditors of the Associated Gas will get a hearing
294 in the 77-B proceeding. In any other sort of proceeding
295 toward collection of that tax, they will not be heard.
296

297 B: Let's assume the Company is perfectly solvent and will
298 pay all its debts.
299
300 W: I can't go along with the idea that you will cooperate
301 with the Government and are ready and able to pay the tax
302 when due.

******

318 W: You know the position the Treasury has taken. We have
319 not intervened--we have not asked the Court yet to be a part
320 to the suit. I have given you what I think are two or three
321 good reasons why that may be the best method of determining
322 tax liability and collecting the tax. That ought to
323 sufficiently demonstrate to you the attitude of the
324 Treasury.
325
326 B: Then I understand the Treasury Department is opposed to
327 our succeeding in the dismissal of that suit?
328
329 W: Yes, the Treasury Department is opposed to seeing that
330 suit knocked out by these amendments to 77-B.
331
332 B: Then I suppose the Department is opposed to seeing 77-B,
333 now pending against us, knocked out at all?
334
335 W: That will develop later.
336
337 B: You are opposed to its being knocked out by legislation
338 by Congress?
339
340 W: That's right.
341
342 B: That is a very interesting position for a Department of
343 the Government to take. I would not have believed it unless
344 you gentlemen told me. I supposed that the Government was
345 not interested in proving a company insolvent. I assumed
346 that the Government was interested in collecting the tax and
347 usually it is considered easier to collect from a solvent
348 company than from one in bankruptcy.
349
350 O: The Treasury is interested in collecting the tax with a
351 minimum of hardship to creditors.
352
353 B: No hardship if you collect it in full.

iii) The Argument. The debate propositions claim that there is no satisfactory explanation for certain actions. The Treasury's position is that there is indeed a good explanation for the actions in question. To support this position they could either (a) make a general argument, akin to a non-constructive proof in mathematics, to the effect that there must

be some (unspecified) good explanation, or (b) show that there is a good explanation by producing and exhibiting a specific satisfactory explanation. They do the latter.

This means that the argument we are primarily interested in, the "means- ends" argument (a) *is an explanation*, (b) *is a sub-part of an (implicit) argument from example* of the form "There is a satisfactory explanation for A, because the following is one: ... ".

The means-ends argument the Treasury side puts forward is quite long and complex. We will first trace its broad outlines, and then progressively fill in its subcomponents.

A good way to start is to consider what argument corresponds to the following two utterances:

(40-43) The object is very simple. We have a suit against you fellows, and we certainly are not going to let a joker be put into some bill which is going to make it impossible for us to go through with this case.

(329-330) Yes, the Treasury Department is opposed to seeing that suit knocked out by these amendments to 77-B.

These utterances tell us that the Treasury Department, TD, has as a goal (wants) to win its suit againt AGEC, or, equivalently*, not to lose its suit against AGEC. They also tell us that in TD's view, the adoption of certain amendments to a bill known as 77-B will cause TD to lose its suit against AGEC.

It is likely that these utterances correspond to an *argument from goals (ends) to sub-goals (subsidiary ends)* as follows:

(G) TD wants (TD wins TD's suit against AGEC).

(C) (X adopts amendments to 77-B) tcause ~(TD wins TD's suit against AGEC).

(Sub-G) TD wants ~(X adopts amendments to 77-B).

The rule of inference which sanctions this argument is:

(R1) If (AGENT wants X) and (Y tcause ~X), then(pf) (AGENT wants ~Y).

There are at least two aspects of the above argument-specification which require immediate comment. First, there is the use of the term "tcause". I want this to be read as "tends to cause", or "creates a causal tendency for". The Treasury

---

* These two states of affairs - TD's winning its suit against AGEC, and TD's not losing its suit against AGEC - are not really equivalent, strictly speaking. In certain circumstances, one may neither win nor loose a suit. However, it seems to me that nothing is lost, and some convenience is gained, by treating them as equivalent here, in context.

undoubtedly does not believe that adoption of the amendments is a necessary causal condition for their losing their suit; they might believe that it is a sufficient causal condition, but only in a pretty narrow (and unspecified) set of conditions (which correspond to the usual crucial but vague and usually implicit "ceteris paribus" clause in most causal claims). The tremendous difficulties involved in spelling out our notion of causal relations are notorious [for a depressing but striking list of unsuccessful approaches, cf. Michael Scriven, "The Logic of Cause", Theory and Decision, Volume 2, 1971, pages 49-66]. In order to be able to proceed with the present analysis, I need to be able to evoke an intuitive notion of cause, as something which "tends to bring about" an effect, without specifying it more than partially. The notion of cause I would like to evoke is that of something (a state, an event, a process, a thing, a relation, a configuration, a thought, or the absence of any of these) which is a sufficient condition for something else (its effect) only in certain circumstances (so a partial or contributory cause is still a cause). According to this notion, causes need not be separated in time or space from their effects, but they are logically separable from their effects, i.e. the connection between cause and effect must be empirical and contingent. This last restriction is perhaps not present in our everyday notion of causality [cf. Scriven, op.cit.] but it is necessary for certain distinctions relevant to the form of means-ends argumentation which will be made below. In our analysis, we will postulate that if X tcauses Y, then if X then(pf) Y, but not vice versa. Secondly, there is the use in (R1) of the "if...then(pf)..." construction. I would like this to be read as "if...then(prima facie)...". I use it rather than the standard logical "if-then" to stress that the inference which this rule sanctions is a *defeasible inference*. Means-ends argumentation is a form of "practical reasoning", and practical reasoning may be contrasted with theoretical reasoning (which certainly includes logical and mathematical reasoning) as follows:

> "An added premise can never invalidate a piece of theoretical reasoning: what follows from a set of premises still follows if the premises are added to [even if the new premise or axiom makes the system inconsistent, the conclusion still follows, trivially - AAA]. But practical reasoning can become invalid from an addition to the stock of premises; for the added premise will express a new end to be achieved, and a policy reasonably inferable from the smaller set of premises - in that it sucures fulfilment of some of the ends then expressed and is not incompatible with any - may be incompatible with the end expressed in the new premise. In this way practical reasoning, unlike theoretical reasoning, is as lawyers say *defeasible*. (page 90)
>
> ...practical reasoning from a set of directives as premises is defeasible by the addition of a premise if its conclusion is incompatible with the fulfilment of that premise, but stands firm if no such premise is added; whereas theoretical reasoning is *never* defeasible by the addition of a premise. (page 115)
>
> ...defeasibility is [also - AAA] a feature of reasonings that relate to *efficient* causality. ... Because of interference and prevention, true causal laws do not state what *de facto* always happens, but only what happens *if* nothing interferes - and that is quite a different matter." (pages 92-93) [Peter Geach, 'Teleological Explanation', in Stephan Korner, ed., "Explanation", Yale University Press, New Haven, 1975]

(R1) would certainly be invalid if it was taken to sanction an undefeasible inference. If it were so taken, it would give irrevocable sanction to an argument such as "John wants to knock down his wall; if an atom bomb were dropped on his home town it would cause his wall to be knocked down; therefore, John want an atom bomb dropped on his home town". But in fact it only gives a provisional sanction to such inferences, and explicitly acknowledges that they are defeasible if other relevant considerations (such as, in our example, John's not wanting to die, and his belief that if an atom bomb were dropped on his home town he would die) were taken into account. So (R1) sanctions a conclusion about an agent's goals only on the condition that no extraneous goals are taken into account; the "if...then(pf).." stresses this restriction, and signals an awareness of the fact that other goals and beliefs may invalidate the conclusion of the inference.

This leads to an important general point: In this section we are studying arguments about means, goals, values and actions, what might be loosely termed *practical argumentation*. And the rules used in practical argumentation are almost all, if not all, *rules of defeasible inference*. They would only provide us with inferential certainty if we were sure that all relevant considerations had been included in our arguments, and we are never, outside of artificially restricted contexts, sure of this.

> "One can never demonstrate a practical conclusion unless one can predict, with full certainty, all of the consequences of all of the actions open to the agent, and specify the agent's entire basis of action, his wants present and future, and the relative desirability of their objects. The sphere of the practical is necessarily the sphere of the uncertain...
>
> In general, a practical argument is satisfactory if the arguer takes reasonable care to determine the sufficiency of the basis, recognizing that to presume to know the agent's future history, whether the agent be himself or another, is absurd." [David P. Gauthier, "Practical Reasoning", Clarendon Press, Oxford, 1963, pages 48-49]

Furthurmore, *if a practical argument is defeasible, then if it constitutes an explanation, that explanation is itself defeasible*, susceptible to being invalidated by consideration of new relevant facts about the agent and the agent's choice situation. Below, whenever we say that a practical argument constitutes an explanation, we will not mean by that it it constitutes an explanation that is in any way final or complete. Explanations in general are in general only more or less incomplete; and the criteria which determine, in context, their degree of completeness depend on the purposes for which they are sought (cf. [Donald Sherer, "Explanatory Completeness", Philosophy, No 188, Vol 49, April 1974]). And this is true in the case of explanations which rest upon practical arguments, in particular.

Let us continue the analysis of the argument. The argument so far explains the Treasury's sub-goal of preventing amendments to 77-B (SG above), but it does not explain any of the actions mentioned in the debate propositions (DP1)-(DP3). However, Morgenthau's utterance in lines 40-43 is clearly meant as an argument against DP3 (three why-questions have already been made at that stage). For our argument to explain the actions, we would have to expand it so as to take account of the following facts. The body which could have amendments to 77-B adopted is the Senate

Committee. 77-B is a piece of legislation having to do with bankruptcy cases, and the proposed amendments are also pieces of legislation dealing with bankruptcy cases. One way of inducing the Senate Committe to reject the amendments is to convey one's opposition to the amendments to the Committee. One way of conveying one's opposition to the amendments is to appear in person, or have one representative appear in person, before the Committee and oppose (argue against, voice opposition to) the amendments.

A preliminary and very undetailed way of expanding the argument would be as follows:

(Sub-G) TD wants ~(Senate Committee adopts amendments to 77-B).

(C) (TD conveys TD's opposition to the amendments to 77-B to the Senate Committee) tcause ~(Senate Committee adopts amendments to 77-B).

(A) TD conveys TD's opposition to the amendments to 77-B to the Senate Committee.

The rule of inference which sanctions this argument is:

(R2) If (AGENT wants X) and (AGENT-ACTION tcause X) then(pf) AGENT-ACTION.*

This argument explains the Treasury Department's conveying its opposition to the amendments to the Senate Committee, but not, on the face of it, the Treasury Department's sending a representative to argue against the amendments before the Committee. Now (A) is a general action description, which refers to a class or set of actions, members or sub- sets of which may be referred to by more specific action descriptions, just as the general noun "tables" refers to a set or class of objects, members or subsets of which may be referred to by more specific descriptions, such as "John Xylappo's night-table" or "the third table Mary built this year" or "dining-room tables". Some descriptions of more specific actions which would count as instances of (A) are:

---

*This rule could be explicitated so as to comprise two sub-rules:

(R2i) If (AGENT wants X) and (Y tcause X) then(pf) (AGENT wants Y).

(R2ii) If (AGENT wants X), (X is an action of AGENT), and (it is possible that X, i.e. if AGENT can perform the action that corresponds to X), then(pf) (X, i.e. the agent performs that action).

(R2i) is clearly very similar to (R1). (R2ii) is a postulate of rationality (such postulates will be discussed below). People are in some respect rational when they do what they conclude, after deliberation, they want. But they often do not do what they consciously, upon reflection, want: they may act out of habit, or on impulse, "against their better judgment".

(A1) The Treasury Department has a representative appear before the Senate Committee to oppose (argue or voice opposition to) the amendments.

(A2) The Treasury writes down its arguments agains the amendments and sends them to the Senate Committee through the mails.

(A3) The Treasury phones the members of the Senate Committee and tells each member that it is opposed to the amendments, and why it is opposed.

.....

Telling the Committee one is opposed to amendments face to face, by writing to them, by phoning them, etc., are all "ways to" convey one's opposition. But if performing a specific action An is a "way to" perform a more general action A (to use a barbaric terminology, if An-ing is a way to A) it does not always follow that performing An counts as, or is an instance of, performing A. The phrase "is a way to" is ambiguous between "is an instance of" and "is a means to". If Jones wants to travel from New York to Chicago, then his flattering Smith so that Smith will give him a ride is, for Jones, a way to travel from New York to Chicago, but is not an instance of his traveling, as taking a plane, train, bus, etc., would be. Perhaps this distinction might be captured by speaking of "ways to" when considering actions which are means to performing (or becoming able to perform) other actions, and speaking of "ways of" when considering actions which count as, or are specific instances of, another more general action. In any case, the essential point is that the relationship between (A) and (A1) is that not of the type which exists between goals and means, or between effects and causes, but is a distinct instantiation relationship. *(A1) is an instance of (A), and (A1) implies ("->") (A), but we cannot say, in our terminology, that (A1) causes (A).*

The above argument explains the general action A; A1 is an instance of A; does the argument therefore explain A1? It seems to me that it does in a weak sense: it explains the performance of at least one instantiation of A. But it does not explain it in a strong sense: that is, it does not by itself explain why one particular instantiation of A was performed rather than another, or some combination of instantiations. That is, it does explain why the Treasury A-s, but not why the Treasury chose a particular way of A-ing. Additional considerations must be brought in to explain this choice - maybe there is some rule saying that arguments must be presented to the Committee in person rather than in writing or by phone. Let us call indeterminacies of this ty *action-instantiation selection indeterminacies.* The action-instantiation indeterminacy we are faced with here could be spelled out as follows:

[1] ((Sub-G) & (C)) -> (A), by rule R2

[2] (Sub-G) & (C)

[3] (A)

[4] (A) <--> (A1 v A2 v A3 v ..... v An)

[5] (A1 v A2 v A3 v ..... v An), from [3] and [4] by modens ponens

[6] though (A1) -> (A1 v A2 v A3 v ..... v A5), it is not the case that (A1 v A2 v A3 v ..... v A5) -> (A1)

[7] therefore it is not the case that (Sub-G & C) -> (A1)

The action-instantiation indeterminacy of (Sub-G)-(A), given the need to explain (A1), is not focussed upon or challenged in the dialogue, because what is important to AGEC is that the Treasury is opposed to the amendments to 77-B, and that it conveys its opposition to the Senate Committee, not *how* it conveys its opposition to the Committee. Given AGEC's lack of interest in how the Treasury conveys its opposition, we can say that the argument we have constructed so far - (G)-(A) -, though in general only adequate to explain (A), but not (A1), is nevertheless functionally adequate in the context of the dialogue to explain (A1).

Though I feel that it would be unwarranted to introduce a very shaky, context-dependent "rule of inference" that would sanction the derivation of A1 from Sub-G, I do feel that the force of the remarks in the previous paragraph could be conveyed by making a second-order statement to the effect that *(Sub-G)-(A), and the larger argument (G)-(A) are explanations of (A), but only partial explanations of (A1).* This would convey our feelings that (G)-(A) increases the intelligibility of (A1), but without making it thoroughly or completely intelligible.* These statements should be recorded, for they are relevant to the Treasury's argument as a whole, which includes the statement "There is a satisfactory explanation for A, because the following is one: ....".

(E1) Arguments of the form (Agent wants X), (Y tcause ~X), (Agent wants ~Y), are explanations of their conclusion.

(E2) Arguments of the form (Agent wants X), (Agent-Action tcause X), (Agent-Action), are explanations of their conclusion.

(E3) If an argument is an explanation of (Agent-Action1), then in conjunction with a statement of the form (Agent-Action2 is an instance of Agent-Action1), it is a partial explanation of (Agent-Action2).

It would also be useful at this point to stress that the explanation relation is such that one can have chains of explanations, but that partiality of explanation is transmitted forward in chains of explanations:

(E4) If an argument {P1...Pl} is an explanation of Pl, and an argument {Pl...Pn} is an explanation of Pn, then the argument {P1...Pl...Pn} is an explanation of Pn.

(E5) If there are two arguments, {P1...Pl} and {Pl...Pn}, either or both of which are partial explanations of their conclusions (Pl and Pn, respectively), then the argument {P1...Pn} is a partial explanation of Pn.

---

*This idea, as did many other ideas in this paper, emerged in discussion with Jim Levin.

As soon as one becomes aware of *action-instantiation selection indeterminacies*, one is likely to think also of *sub-goal selection indeterminacies*, and *means-selection indeterminacies* in practical argumentation. The former of these involves the instantiation relationship between actions, whereas the latter two involve tcausal relationships between objects, actions, states, events, and so on. Let us describe and compare these.

a) Action-instantiation selection indeterminacies arise when, as in the example above, an agent has several (a set of) ways of performing a general action he wants to perform, and yet chooses only to perform one (or a sub-set) of them. His desire to perform the general action only constitutes a partial explanation of his performing the particular action- instantiation because it leaves the question "Why does he choose to perform this particular action-instantiation rather than that?" unanswered.

b) Sub-goal selection indeterminacies arise when several different actions, states or events would cause the action, state or event desired by the agent to obtain or occur, and yet he only designates one (or a sub- set) of them as consequently desired, as his sub-goal. Again, his goal and the causal relations only consititute a partial explanation of his sub-goal, because it leaves the question "Why does he choose this sub-goal rather than that?" unanswered.

c) Means-selection indeterminacies arise when several different courses of action are open to the agent, each of which would cause his goal action, state or event to obtain or occur, and yet he chooses to pursue only one (or a sub-set) of them. By themselves, his goal and the causal relationships between the means and his goal only constitute a partial explanation of his pursuing one (or a sub-set) of the courses of action, for it leaves the question "Why does he choose this means of attaining his goal rather than that?" unanswered.

(a)-(c) presuppose choice, and a selective decision. For all of these indeterminacies to exist, the agent must (a) believe that he has more than one option open to him, and (b) select among the options (not adopt all of them). In the argument (Sub-G)-(A) above, it is postulated that the agent thought that he had only one means at his disposal; he was not faced with a choice, and consequently, (Sub-G)-(A) does not suffer from means- selection indeterminacy, and provides an explanation, not a partial one.

If the agent has a choice (between action-instantiations, states/events/actions which causally bring about his goal), and makes a selective decision, then the argument explaining his selective decision will contain several statements representing those options: several statements of the form (An is an instance of A) or of the form (Y tcause X), as the case may be. It will also contain statements to the effect that he performed fewer actions or adopted fewer sub-goals than he could have.

An argument of this type does not necessarily constitute a partial explanation, however. It will only constitute a partial explanation *it it lacks a comparative evaluation statement* to the effect that the actions the agent selectively chose to perform, or the states he selectively chose as sub-goal states, are the best of those he had to choose from. In the latter half of this section, we will construct a goal-to-action argument in

which several means are mentioned, but which also comprises a comparative evaluation showing that the means chosen was the best one; it constitutes a full, not a partial, explanation.

We should state the above remarks as a (very loosely expressed!) *meta- rule of explanation*, (Meta-E1). We will also modify our rules (R1) and (R2), and the rules of explanation which correspond to them, (E1) and (E2), so as to capture these generalities. (Below, curly brackets are used to signify optionality.)

(Meta-E1) If an argument represents a situation where the agent has only one option, and where he adopts that option, then it constitutes an explanation of its conclusion.

If an argument represents a choice situation where there are several options, and either represents an adoption of all the options, or represents an adoption of less than all the options and includes a comparative evaluation of the options which shows that the options selected are the best of the set of options, then it constitutes an explanation of of its conclusions.

If an argument represents a choice situation where there are several options, and an adoption of less than the full set of options, but lacks a comparative evaluation of the options which shows the subset of options selected to be the best of the set of options, then it is only a partial explanation of its conclusion(s).

(R1) If (Agent wants A), (B tcause ~A){.....(Z tcause ~A)}, {(B{...X} is/are the worst of the set B...Z)}, then(pf) (Agent wants ~B){...(Agent wants ~Z)}. (Note the reversal here of "best" and "worst", due to the fact that the conclusion is of the form "Agent wants ~X".)

(E1) Arguments of the form (Agent wants A), (B tcause ~A), (Agent wants ~B), are explanations of their conclusion.

Arguments of the form (Agent wants A), (B tcause ~A)....(Z tcause ~A), (B...X are the worst of the set B...Z), (Agent wants ~B)...(Agent wants ~X), are explanations of their conclusion(s).

Arguments of the form (Agent wants A), (B tcause ~A)...(Z tcause ~A), (Agent wants B)...(Agent wants X), are only partial explanations of their conclusions.

(R2) If (Agent wants A), (Agent-ActionA tcause A){,...(Agent-ActionZ tcause A)}, {(Agent-ActionA...Agent-ActionX are the best of the set Agent-ActionA...Agent-ActionZ),} then(pf) (Agent-ActionA){...(Agent-ActionZ)}.

(E2) Arguments of the form (Agent wants A), (Agent-Action tcause A), (Agent-Action) are explanations of their conclusions.

Arguments of the form (Agent wants A), (Agent-ActionA tcause A)....(Agent- ActionZ tcause A), (Agent-ActionA...Agent-ActionX are the

best of the set Agent-ActionA...Agent-ActionZ), (Agent-ActionA)...(Agent-ActionX) are explanations of their conclusions.

Arguments of the form (Agent wants A), (Agent-ActionA tcause A)....(Agent- ActionZ tcause A), (Agent-ActionA)...(Agent-ActionX) are only partial explanations of their conclusion(s).

At this point we should ask whether the argument {(G)-}(Sub-G)-(A) explains all of the actions mentioned in the debate propositions (DP1)- (DP3). The argument explains (A) and partially explains (A1); let us compare these to the three actions mentioned in (DP1)-(DP3), which we will call (DA1), (DA2), and (DA3).

(A) TD conveys TD's opposition to the amendments to 77-B to the Senate Commitee.

(A1) TD has a representative appear before the Senate Committee to oppose (argue or voice opposition to) the amendments to 77-B.

(DA1) A representative of TD appears before the Senate Committee.

(DA2) A representative of TD appears before the Senate Committee to oppose legislation having to do with bankruptcy cases.

(DA3) A representative of TD appears before the Senate Committee to oppose legislation having to do with bankruptcy cases on the ground that it would relieve AGEC of 77-B proceedings and be beneficial to AGEC.

Given that we have a partial explantation of (A1), we also have a partial explanation of (DA2), because (A1), in conjunction with the postualted fact that the amendments to 77-B are legislation having to do with bankruptcy cases, and the semantic (presuppositional) rule that if Agent1 has Agent2 do X, then Agent2 does X, logically implies (DA2). There is a general rule about explanation involved in our reasoning here, viz.:

(E6) If an argument, ARG, (partially) explains X, and X -> Y, then ARG (partially) explains Y.

For example, if we have an explanation for why our steak burned, and our steak was the third article we bought in the market this morning, so that {(Our steak burned) & (Our steak is the third article we bought in the market this morning)} -> (The third article we bought in the market this morning burned), then we have an explanation for the fact that the third article we bought in the market this morning burned. And by the same rule, we also have a partial explanation for (DA1), since we have a partial explanation for (DA2) and (DA2) implies (DA1).

However, we do not have a (partial) explanation for (DA3), for neither (A), (A1), nor (DA2) imply (DA3). (DA3) is a complicated action: a physical act done with a purpose and with certain explanatory and/or justificatory reasons. A representative of TD appears before the Senate Committee in order to oppose the amendments to 77-B (legislation having to do with bankruptcy cases), opposes them, and has or gives as his reasons for opposing them (a) that they would relieve AGEC of 77-B proceedings and

(b) that they would be beneficial to AGEC. I say "have or give" because of the ambiguity mentioned in footnote * above: it is not clear whether the representative opposed the amendments for reasons which he did not express to the Committee, or whether he in addition told the Committee about reasons he had which explained and/or justified his opposition. If we are to specify what changes in the argument (G)-(A) are necessary for it to explain (DA3), we must ask ourselves what kind of considerations would explain someone opposing something for certain reasons, i.e. having or giving certain reasons for his opposing something. This requires a digression, a short examination of what reasons and reason-giving are.

The term "reason" is ambiguous in several ways, two of which are directly relevant here:

> "...we first need to note an ambiguity in the term "reason". In the first usage (which I shall term reason1), a reason is taken to be a statement, proposition or sentential clause which makes an assertion or describes a particular state of affairs. In this sense, providing a reason for acting amounts to asserting some proposition p, which constitutes the reason.(4) [footnote 4: For an example of this usage, see A.I.Melden, *Free Action* (London: Routledge and Kegan Paul, Ltd., 1961), especially pp. 160- 167. "If someone stops his car and is asked "Why?", the statement "There is a restaurant nearby" is not to be understood merely as a bit of information, but as a reason for the action." (pg. 160)...] In the second usage (which I shall term reason2), a reason is the belief, desire, want or other psychological state of the agent. In this sense, if p is a proposition or sentential clause describing a particular state of affairs, then the reason in question is the agent's belief that p, his desire that p, his wanting that p, etc.. Examples can be found of both senses of "reason", and failure to note tha ambiguity may result in conclusions which are true of "reason" in one sense but not true in the other." (pg. 80)
>
> "...it is important to note a possible ambiguity in the term "reason". The term can cover not only those reasons we hold to be truly explanatory, but also the kind of reasons offered in *justification* of a particular action. Reasons of the latter type do not, strictly speaking, explain the action. Rather, they are reasons the agent gives in justifying his behavior, whatever the explanation may be. ... Since reasons may be offered in justification of an action - reasons which are not explanatory - not all reasons will be candidates for entrance into causal explanations." (pg. 88) [Ruth Macklin, "Reasons Versus Causes in Explanation of Action", Philosophy and Phenomenological Research, Volume 33, No 1, September 1972]

When we explain an agent's actions, we give *his* reasons for acting, and these are reasons2 - statements about his beliefs, desires, motives, intentions and so on. But we often also give reasons1, statements about states of affairs which the agent was not necessarily aware of (either consciously, or unconsciously). For example: "Jones made Smith angry, because he wanted to drink some hot tea (reason2), and poured it into Smith's prize cyrstal glass. Jones thought the glass was heat-resistant (reason2), but it wasn't (reason1), and broke."*

In our reconstruction of TD's explanatory arguments, *we suppose that all the statements in those arguments are reasons2.* This is a very reasonable assumption,

given that TD is explaining its own actions. So all of the statements in the arguments we construct could be thought of as being embedded in an implicit "TD believes that ....." phrase.

Reason-explanations for actions appeal to (but need not explicitly mention) a whole *system* of beliefs and motives, wants, attitudes, intentions and so on [R.C.Solomon, "Reasons as Causal Explanations", Philosophy and Phenomenological Research, Volume 34, No 3, March 1974, pg.416 and pg.423]. They are understood, and are satisfactory as explanations, because they reconstruct the result of the agent's deliberation which lead to his action, and allow us to "put ourselves in the agent's place" [ibid., pg.418]. This has three important consequences: (a) reason- explanations are appropriate explanations only of considered, reflective action, (b) involve the attribution of some degree of rationality to the agent whose action is being explained, and (c) consist of a reconstruction, not of the psychological process of deliberation the agent went through before perfoming the act, but of the piece of practical reasoning the agent's deliberation enabled him to produce. To elaborate the first point: we do not give reason-explanations of actions which we regard as purely reflexive, impulsive or habitual. We do give such explanations for actions about which the agent deliberated prior to his action, and came to make some practical judgment about the action. In other words, we give reason-explanations only for considered, reasoned acts. It is important to note, however,

> "Deliberate action, and only deliberate action, is *reasoned* action, action performed *for* reasons. Reasoned action is not to be confused with reasonable action. No doubt the agent supposes his deliberate, reasoned actions to be reasonable. But he may refuse to consider important features of the situation in which he acts. He may ignore some of the consequences which is action will have. However reasoned his actions, he may not be a reasonable man. Thus reasoned action may be unreasonable. Conversely, reasonable actions may be performed quite without prior thought, even on impulse. An agent may have good reason to do what he does, although he does not consider this in determining what to do. Thus reasonable actions may not be reasoned." [David P. Gauthier, "Practical Reasoning", Clarendon Press, Oxford, 1963, pgs.26-27]

Since reason-explanations are explanations of reasoned action, they presuppose some degree of rationality on the part of the agent.

> "At least one of the presuppositions in any general account of human behavior is an assumption of rationality on the part of the agent. We tend to hold this assumption fixed, unless there are indications to the contrary. Minimally, this assumption entails that an agent's preferences are transitive and asymmetrical; that if an agent prefers x to y, ceteris paribus, he chooses x over y; that in the absence of intervening factors, if an agent decides to do x he does x; that an agent acts in such a manner as

---

*This example might be felt to be weak and controversial, in that it constitutes an unintentional action. It should be possible, however, to generate an example of an intentional action that was facilitated by conditions of which the agent was unaware.

to maximize expected utility. In a rough and ready way, we say that these (and other * - cf. the extended footnote on the next pages) factors

---

*The following principles of rationality have been specified by Kai Nielsen ["Principles of Rationality", Philosophical Papers, Volume 3, October 1974, pgs.57-58]:

1. Relevant evidence or considerations are, ceteris paribus, not to be ignored in the forming or holding of beliefs.

2. Objectivity is to be maintained or at least striven for. Relevant evidence and considerations are, ceteris paribus, to be duly taken into account or at least conscientiously sought.

3. Beliefs are ceteris paribus, to be striven for, for which it is known that there are good grounds for believing that they do not involve inconsistencies or contradictions.

4. Beliefs are, ceteris paribus, to be striven for, for which it is known that there are good grounds for believing they do not involve incoherencies.

5. The most efficient and effective means are to be taken, ceteris paribus to achieve one's ends.

6. If one has several compatible ends, one, ceteris paribus, is to take the means which will, as far as one can ascertain, most likely enable one to realize the greatest number of one's ends.

7. Of two ends, equally desired and equal in all other relevant respects, one is, ceteris paribus, to choose the end which one has good grounds for believing has the higher probability of being achievable.

8. If there are (as far as one can ascertain) the same probabilities in two plans of action, which secure entirely different ends, that plan of action is, ceteris paribus, to be chosen which secures ends at least one of which is preferred to one of those secured by the other plan.

9. If one is unclear about what one's ends are or what they involve or how they are to be achieved, then, ceteris paribus, a postponement is to be made in making a choice among plans of action to secure those ends.

10. Those ends, which, from a dispassionate and informed point of view, one values absolutely higher than one's other ends, are the ends which, ceteris paribus, are to be achieved. A rational agent will, ceteris paribus, seek plans of action will satisfy those ends; and plans to satisfy his other ends will be adopted only in so far as they are compatible with the satisfaction of those ends he or she values most highly.

11. Ceteris paribus, one is to engage in prudent maximizing, i.e. an agent is to maximize the satisfaction of his or her interests.

12. Rational beliefs are beliefs for which one has or could readily come to have good evidence. ...

13. Rational beliefs are critical beliefs; that is to say, they are beliefs which are held open to refutation or modification by experience.

14. Rational beliefs are beliefs which are held in such a way that those holding them will not resist attempts critically to consider their assumptions, implications and relations to other beliefs. They will be beliefs which are open to reflective critical inspection.

15. A rational person's actions, ceteris paribus, will generally be in accordance with his or her rational beliefs.

constitute our attribution of rationality to the agent. If there are indications that one or more of these factors is not present, then we introduce special factors to explain an action which fails to conform to the rational pattern. It is often alleged that these special factors can be viewed as causal factors since they can in now way be brought in as the agent's reasons for acting. But where the action can be explained in terms of the agent's rational calculation, here it is appropriate to give reasons." (pg. 84)

"...we attribute aims, goals, and purposes to human agents.... But the only way in which we are able to employ these aims, goals, and purposes in our reason explanations of human action is by holding fixed the assumption of rationality. That is, the explanatory force of a given reason depends, among other things, on the presupposition of rationality on the part of the agent. Where rationality in an agent breaks down, we need to introduce another set of factors (probably also in terms of internal psychological states) in order to explain the failure of the usual standing conditions to obtain." (pg. 85) [Ruth Macklin, op.cit., September, 1972]

Lastly, a reason-explanation is not a reconstruction or simulation of the psychological process of deliberation the agent in fact went before he acted. Deliberation is not effected by practical reasoning, or by any formal pattern of reasoning whatsoever. "To speak of deliberation as a type of reasoning is to point to the fact that, as a result of successful deliberation, one can produce a piece of reasoning, an ordered argument, leading from a starting point ... to a conclusion - an action to be done." [David P. Gauther, op.cit., 1963, pg.26] A reason-explanation reconstructs the argument the agent's deliberation enabled him to produce; in this argument, the agent's steps leading to his resolution of his practical problem (What should I do?) are formally set out. **

**This point is an instance of the general principle that logic and argumentation generally do not correspond directly to the psychological procedures which generate them. This principle has been put as follows in the case of logic. "There is a use of the term "inference" in accordance with which *an inference* is a set of propositions one member of which is a logical consequence of the rest taken jointly. If one were to ask why a study of inference in this sense was important, I would answer that it was important at least in order to understand the concept of a proof, and in order to develop a systematic way of telling whether something was indeed a proof. But I would certainly *not* say that the study of inference was important in order to discover the nature of the procedures we employ or of the mental events which might in fact occur when we come to believe this proposition or when we come to disbelieve that one. The reasons I would not say anything of this sort are these. First, we do not always think and act in accordance with deductive norms. When one believes one proposition because he believes another, his coming to believe the one just may not be explicable in terms of the fact that it is a logical consequence of the other, and a very good reason for this might be that it simply is *not* a logical consequence of the other. Secondly, even in those cases where our thinking is deductively sound, it is not always true that the sequence of our thoughts is isomorphic to the inference patterns which constrain us." [Stephen E. Norris, "The Intelligibility of Practical Reasoning", American Philosophical Quarterly, Volume 12, No 1, January 1975, pgs.77-78]

The distinction between explanatory reasons and justificatory reasons is an important one. One can give explanatory reasons which are not justificatory - "I fired a shot through the window because I wanted to kill Mr. Jones" - just as one can give reasons to justify an action without thereby explaining it - "Of course I voted in the Presidential election; it was my duty as a citizen to do so", uttered by someone who avowedly never gives a thought to his duty, and only voted because he hated one of the candidates. Nevertheless, however important the distinction between the explanatory and the justificatory functions of reasons may be, many reasons fulfill both functions simultaneously (consider "I refuse to answer on the grounds that it might incriminate me", uttered by a witness at a trial). One may explain one's action by giving one's reasons for acting, but if one's reasons are *the* reasons, or *good* reasons, one at the same time justifies one's action. The question of what makes an explanatory reason a justificatory reason is a complicated one, but a very rough preliminary answer to it might be that when reasons involve beliefs (especially evaluative beliefs), goals, desires, intentions and plans which conform to socially accepted values, ideals, norms, and moral, legal, political or religious rules, principles and ideologies, they tend to be justificatory as well as explanatory.

Let us return from our digression to the problem at hand. The representative of the Treasury Department appeared before the Senate Committee and opposed the amendments. He opposed them on certain grounds. To do something or to have an attitude on grounds is certainly to have and to (at least) be ready to express reasons for doing that thing or having that attitude. As was mentioned twice above, it is unclear whether the representative just had certain reasons, or whether he expressed them before the Committee.

Let us first consider the (simpler) case where the representative just had these reasons for his opposition. It might be thought that since argument (G)-(A) is a partial explantation of (DA2), it is also a partial explanation for (DA2) being done for reason R, where R is a proposition included in (G)-(A). One might think that this conclusion is warranted by the following rule of explanation:

(E7) If the argument {P1 ... Pn} constitutes a (partial) explantion for an action A, then it is a (partial) explanation of Ar, Ar = A is done for reason2* R, where R is a member of the ordered set {P1 ... Pn}, or an element of one of the members.**

(E4) seems to be an operative rule. Suppose I wrote a letter to Jones. There is a partial explanation of this act which goes as follows: I wanted Jones to know that his sister had graduated; my conveying that information to Jones would cause him to know it; I wanted to convey the information; writing the letter was one way of conveying the

---

*Cf. the distinction between reasons1 and reasons2 above. This rule requires that R be a conscious belief of the agent of the action A.

**This last phrase means that if, for example, the argument contains a proposition of the form A -> B, then R could be either A or B as well as, of course, the whole proposition A -> B.

information; I wrote the letter. This very same explanation would also be a partial explanation of my writing to Jones because (for the reason that) I wanted him to know that his sister had graduated or of my writing to Jones because that was one way of conveying the information that his sister had graduated.

The use of (E4) seems to be the right approach. But two important matters must be dealt with before it can be applied to solve the problems posed by (DA3).

First of all, two points, (a) and (b). (a) (G)-(A) is a partial explanation of (A1), "The Treasury Department has a representative (of TD) appear before the Senate Committee to oppose the amendments.". (A1) implies (DA2), "A representative of TD appears before the Senate Commitee to oppose legislation having to do with bankruptcy cases". (It implies it by virtue of (a) the fact that "the amendments" = "legislation having to do with bankruptcy cases", and (b) the semantic rule that "Agent1 has Agent2 do A" presupposes "Agent2 does A".) So, by (E3), (G)-(A) is a partial explanation of (DA2). Rule (E4) warrants saying that (G)-(A) is a partial explanation for the Treasury Department's doing the action described in (A1) for the reason that if the Committee adopts the amendments the TD would not win its suit against AGEC ((C) above). But this is insufficient in at least two ways. (a) (C) is not prima facie equivalent to the first reason mentioned in (DA3), which can be interpreted as "if the Committee adopts the amendments then AGEC is relieved of 77-B proceedings". We have to provide an argument containing additional premises to show that if the TD does not win its suit against AGEC, AGEC is relieved of 77-B proceedings. The question then arises of how we would want to make use of this implication. We might be tempted to suppose that if an explanation provides a (partial) explanation for an action performed for a reason, then it provides a (partial) explanation for that action being performed for any reason implied by that reason, i.e. we might try to use a rule of explanation such as the following:

(XE8) If an argument is a (partial) explanation for an action being done for a reason2 R1, and R1 -> R2, and the agent of the action consciously believes that R1 -> R2 * , then the argument is a (partial) explanation for the action being done for reason2 R2.

---

*This condition would be crucial for the soundness of this hypothetical rule. If it were not imposed, there would be counter-examples such as the following. Suppose that I light a match in a warehouse for the reason2 that if I light a match, then I can smoke. Suppose that that reason implies that if I light a match, then the warehouse will blow up (because there are explosives stacked all around). I would not want to say that I light a match for the reason that doing so would blow up the warehouse.

.Note also that since (XE8) states that R2 is a reason2, it assumes that if someone believes X and also believes X -> Y, they will belive Y. This is itself often a questionable assumption; cf. [Nicholas Rescher, "Epistemic Modality: The Problem of a Logical Theory of Belief Statements", in Nicholas Rescher, "Topics in Philosophical Logic", D.Reidel, Dordrecht- Holland, 1968].

But a moment's reflection shows that (E5) is invalid as it stands if it justifies such moves. Suppose that there is some explanation of my staying out of the California coastal waters because (for the reason that) I believe some of the fish therein are dangerous. If I believe that some of the fish in the Californian coastal waters are dangerous, I believe, let's say, that some fish in some parts of the oceans are dangerous. But I do not stay out of the Californian coastal waters because I believe only that some fish in some parts of the ocean are dangerous. It is my specific, not my general belief, which is an explainable reason for my action. So we cannot use (E5). Rather, we must incorporate the statement that TD does not want AGEC relieved of 77-B proceedings *in the body of the explanatory argument*, and use (E4). This incorporation is quite easily done, as follows:

(G2) TD wants (TD wins TD's suit against AGEC).

(C) (X relieves AGEC of 77-B proceedings) tcause ~(TD wins TD's suit against AGEC).

(Sub-G) TD wants ~(X relieves AGEC of 77-B proceedings). by (R1)

(b) If (G)-(A) is a partial explanation for the Treasury Department's having done the action described in (A1) for a certain reason2, that does not prima facie mean that (G)-(A) constitutes a partial explanation for the Treasury Department's representative's having done the action described in (DA2) for that reason2. To show that we do have a partial explanation for the representative's acting for a reason2, which is presumably the same as that which the TD had for sending him in front of the Committee, we would have to bring in a highly questionable additional premise, something like: "if an organisation has a representative do A for reason2 R, then the representative does A for reason2 R.". This is a very real problem, but not one which really emerges in the dialogue.

Secondly, there are two reasons involved: (1) the legislation would relieve AGEC of 77-B proceedings and (2) the legislation would be beneficial to AGEC. As we have just seen, (1) is a consequence of the legislation, if passed, causing the TD to lose it's suit against AGEC, and so can be explained as a reason, if it is incorporated in the body of the explanatory argument. But (2) is nowhere to be found in (G)- (A). (1) does imply (2), however, with the help of some plausible premises to the effect that having proceedings against one is harmful and not having proceedings is beneficial. So it might be thought that we should include the derivation of (2) from (1) in the body of the argument, and thus be able to explain the TD's opposition for reason2 (2) by virtue of (E4). However, from my understanding of the dialogue I do not feel that this step would be justified: *(2) is a reason that is mentioned by AGEC as a reason for the TD's opposition, but is at no point accepted or explained as a reason by the Treasury Department.* Indeed, it would be surprising if (2) were accepted by the Treasury, for it would mean that the TD was voluntarily being unfair to AGEC, something which it explicitly does not want to do (line 63 - we will consider this goal below). The TD is not opposed in general to anything that is beneficial to AGEC; rather, it is opposed to one specific action or event - AGEC's being relieved of 77-B proceedings - which happens to be beneficial to AGEC. Here again we have an example which shows the invalidity of the rule (E5) which we contemplated accepting above. Since (2) is a reason for which an explanation is required but for which no explanation is provided, we will not modify our representation of the TD's argument so as to make it an explanation of (2).

Let us now consider the more complicated case in which the representative expressed explanatory and/or justificatory reasons for his opposition. An explanation of this action would involve a whole new set of means-ends reasoning about why, given the representative's goals, which presumably would be preventing the adoption of the amendments by the Senate Committee, he chose to express the reasons he did. This would involve making many complex and debatable statements about persuasiveness of arguments or reasons in general, and the persuasiveness of certain arguments or reasons for the Senate Committee in particular. It would require the knowledge of many forensic, historical and legal facts. Since (a) this would be difficult and lengthy, (b) many of the theoretical problems associated with means-ends argumentation are considered in this section anyway, and (c) the question of whether the representative did indeed voice the reasons, and why he did, is not a major issue in the dialogue, I have decided, with Jim Levin's assent, not to attempt such an arduous explanation here.

At this point in our discussion we have examined two goal-to-sub-goal arguments and one goal-to-means argument. Before we complete the first half of this section by specifying our reconstruction of the TD's argument up until that point, we should increase our stock of rules and argument- forms by considering an argument which we can extract from lines 151-157. We will first display the argument, and then discuss the rules which sanction its inferential steps.

(G1) TD wants (TD does TD's job fairly and honestly). (line 153)

(Inst1) (TD protects the people of the US) is an instance of (TD does TD's job fairly and honestly). {from some theory of the role of the TD, and some theory of justice, fairness and honesty, both unspecified}

(Sub-G1) TD wants (TD protects the people of the US) {from (G1),(Inst1) by (R3)}

(Explan1) The argument (G)-(Sub-G1) is a partial explanation of (Sub-G1). {by (E8)}

(Inst2) (x)(if x is a case, then the people of the US can try x fairly) is an instance of (TD protects the people of the US).

(Sub-G2) (TD wants ((x)(if x is a case, then the people of the US can try x fairly)). {from (Sup-G1),(Inst2) by (R4) and the semantics of the verb "to protect"}

(Explan2) The argument (Sub-G1),(C1),(Sub-G2) is an explanation of (Sub-G2). {by (E9)}

(I2)

If (Ex)[x is legislation, and x tcause (Ey)(y is a case, and the people of the US cannot try y fairly) *]

then ~((x)(if x is a case, then the people of the US can try x fairly)).

(Sub-G3) TD wants ~(Ex)(x is legislation, and x tcause (Ey)(y is a case, and the people of the US cannot try y fairly). {from (Sub-G2),(I2) by (R5)}

(Explan3) The argument (Sub-G2),(I2),(Sub-G3) is an explanation of (Sub-G3). {by (E10)}

(I3) (x)(if x = amendments to 77-B, then x is (a piece of) legislation)

(I4) (Ex)(x = the amendments to 77-B) -> (Ex)(x is (a piece of) legislation) {from (I3)}

(I5) (x)(if x = TD's suit against AGEC, then x is a case)

(I6) (x)(if (x = the amendments to 77-B) and (The Senate Committee adopts x), then (x tcause (the people of the US cannot try TD's suit against AGEC fairly)

(I7) (Ex)((x = the amendments to 77-B) and (The Senate Committee adopts x)) -> (Ex)(x is (a piece of) legislation, and (Ey)(y is a case, and the people of the US cannot try y fairly)). {from (I3)-(I6)}

(Sub-G4) (TD wants ~(Ex)((x = the amendments to 77-B) and (The Senate Committee adopts x)). {from (Sup-G3),(I7) by (R5)}

This may also be phrased more simply as (TD wants ~(The Senate Committee adopts the amendments to 77-B)), where the force of (Ex)(x = the amendments to 77-B) is conveyed by existential presupposition.

(Explan4) The argument (Sub-G3),(I7),(Sub-G4) is an explanation of (Sub-G4). {by (E10)}

(Explan5) The above argument is an explanation of (Sub-G4). {from (Explan1)-(Explan4) by repeated application of (E4)}

This is an extended goal-to-sub-goal argument, with component sub- arguments of different types. To the extent to which (G1) is a socially condoned goal, it may provide not only an explanation, but also a justification of (Sub-G4). The rules it uses are discussed below.

The argument from (G1) to (Sub-G1) is a derivation of a sub-goal of the type (Agent wants (Agent-ActionB)) from a goal of the form (Agent wants (Agent-ActionA)) and an action-instantiation statement, (Agent-ActionB is an instance of Agent-ActionA). It is sanctioned by (R3). Arguments of this general type can be either full or partial explanations of their conclusions - depending on whether they represent a selective choice between several action-instantiations or not, and whether they contain comparative evaluations of the action-instantiations or not (cf. (Meta- E1) above). We specify this fact in rule (E8):

---

* This expression, of the form (Ey)(A and ~B), is of course equivalent to the negation of an expression of the form (Ay)(A -> B), and is thus the simple negation of what TD is said to want in (Sub-G2).

(R3) If (Agent wants Agent-ActionA), (Agent-ActionB is an instance of Agent-ActionA) {(Agent-ActionC is an instance of Agent-ActionA), ..., (Agent-ActionZ is an instance of Agent-ActionA)}.

then(pf) (Agent-ActionB) {,(Agent-ActionC) {(Agent-Action4) {, ... {,(Agent-ActionZ)}}}}

(E8) Arguments of the form (Agent want Agent-ActionA), (Agent-ActionB is an instance of Agent-ActionA) , (Agent-ActionB) are explanations of their conclusion.

Arguments of the form (Agent wants Agent-ActionA), (Agent-ActionB is an instance of Agent-ActionA).....(Agent-ActionZ is an instance of Agent-ActionA), (Agent-ActionB.....Agent-ActionX are the best of the set Agent-ActionB....Agent-ActionZ), (Agent-ActionB).....(Agent-ActionX) are explanations of their conclusions.

Arguments of the form (Agent wants Agent-ActionA), (Agent-ActionB is an instance of Agent-ActionA).....(Agent-ActionZ is an instance of Agent-ActionZ), (Agent-ActionB).....(Agent-ActionX) are only partial explanations of their conclusions.

The argument from (Sub-G1) and (I1) to (Sub-G2) is sanctioned by (R4). Roughly, the idea behind (R4) is that if one wishes a state of affairs to obtain, and if S1 obtains then another state S2 obtains, because S2 is an instance of S1, then one wants S2 to obtain.

One state, S2, is an instance of another state S1, roughly, if S1 -> S2, but it is not the case that S1 tcause S2. Suppose S1 was a state of affairs in which Saudia Arabia rules the world, and that S2 is a state of affairs in which Saudia Arabia rules France. We would say that S2 is an instance of the state S1. S1 implies S2, because France is part of the world. But we could not say that S1 caused S2.

There is an important difference between the instantiation relationship between states and the instantiation relationship between actions. As we have seen, an action is equivalent to the disjunction of its instantiations. But a state is equivalent to the conjunction of its instances. Because of this, we do not have the kind of indeterminacy in the case of arguments involving state-instantiation as we do in the case of arguments involving action-instantiation.

These considerations raise a host of interesting and intricate questions about the relationship between actions and states, the relationship of both of these to causality, implication and instantiation, and the distinctions between causality, implication and instantiation. Pressed for time, however, we must ride roughshod over these problems, and go on forthwith to specify (R4) and the associated rule of explanation (E9).

(R4) If (Agent wants State1), and (State2 is an instance of State1), then(pf) (Agent wants State2).

(E9) Arguments of the form (Agent wants State1), (State2 is an instance of State1), (Agent wants State2) are explanations of their conclusions.

The reader may be puzzled by the fact that in the argument sanctioned by (R3), we interpreted "(TD protects the people of the US)" as referring to an action, whereas in applying (R4), we are interpreting that same expression as referring to a state, one in which the people are protected. We feel we can do this because the semantics of the verb "to protect" (and of many other, though not all verbs) is such that for X to protect Y, X must be successful, i.e. effectively produce a state in which Y is protected. So the expression "TD protects the people of the US" designates both an action and a state.

The action in question does indeed cause the state of affairs in which the people are protected; but the state of affairs in which they are protected does not cause their being able to try a case fairly. Rather, the state of affairs in which they are protected (S1) implies a state of affairs in which they are able to try a case fairly (S2); S2 is a necessary condition or instance of S1.

The arguments (Sub-G2),(I2),(Sub-G3) and (Sub-G3),(I7),(Sub-G4) are sanctioned by (R5). The idea behind (R5) is that if one desires a goal state S1, and if S2 implies ~S1, so that ~S2 is a necessary condition for S1, then one desires ~S2. Here again, it seems to me, there is no indeterminacy.

(R5) If (Agent wants State1), and (State2 -> ~State1), then(pf) (Agent wants ~State2).

(E10) Arguments of the form (Agent wants State1),(State2 -> ~State1), (Agent wants ~State2) are explanations of their conclusions.

At this point it would undoubtedly be helpful to lay out our reconstruction to date of the TD's argument. It consists of the last argument above (the goal-to-sub-goal argument) appended to the goal-to- action argument directly below. The two arguments link up through the goal (TD wants ~(The Senate Committee adopts the amendments to 77-B)), which is (Sub-G4) above and (Sub-G1) below. There are thus two goals of the TD which explain it.

(G) TD wants (TD wins TD's suit against AGEC).

(C1) (The Senate Committee adopts the amendments to 77-B) tcause ~(TD wins TD's suit against AGEC).

(C2) (The Senate Commitee adopts the amendments to 77-B) tcause (X relieves AGEC of 77-B proceedings).

(C3) (X relieves AGEC of 77-B proceedings) tcause ~(TD wins TD's suit against AGEC).

(Sub-G1) TD wants ~(The Senate Committee adopts the amendments to 77-B). {from (G),(C1) by (R1)} {also from (C1),(C2),(C3), by transitivity of the tcause relation and (R1)}

(Sub-G2) TD wants ~(X relieves AGEC of 77-B proceedings). {from (G),(C3) by (R1)}

(C4) (TD conveys TD's opposition to the amendments to 77-B to the Senate Committee) tcause ~(The Senate Committee adopts the amendments to 77-B).

(A1) (TD conveys TD's opposition to the amendments to 77-B to the Senate Committee). {from (Sub-G1),(C4), by (R2)}

(I1) (for all x)(if x is the amendments to 77-B, then x is (a piece of) legislation having to do with bankruptcy cases).

(A1') (TD conveys TD's opposition to (a piece of) legislation having to do with bankruptcy cases to the Senate Committee). {from (A1),(I1), by substitution}

(Expln1) (The argument (G)-(A1') is an explanation of (A1'). {by (E1)}

(Inst1) (The TD has a representative of TD appear before the Senate Committee to oppose a piece of legislation having to do with bankruptcy cases) is an instance of (The TD conveys TD's opposition to a piece of legislation having to do with bankruptcy cases to the Senate Committee).

(A1'.1) (The TD has a representative of TD appear before the Senate Committee to oppose a piece of legislation having to do with bankruptcy cases).

(Expln2) (The argument (G)-(A1') is a partial explanation of (A1'.1)). {from (G)-(A1'),(Inst1) by (E2)}

(Presup1) (The TD has a representative of TD appear before the Senate Committee to oppose a piece of legislation having to do with bankruptcy cases) -> (A representative of TD appears before the Senate Committee to oppose a piece of legislation having to do with bankruptcy cases). {by semantic presupposition}

(A1'.1.1) (A representative of TD appears before the Senate Committee to oppose a piece of legislation having to do with bankruptcy cases). {from (A1'.1),(Presup1) by modens ponens}

(Explan3) (The argument (G)-(A1') is a partial explanation of (A1'.1.1). {from (Explan2),(Presup1) by (E3)}

(A1'.1+R) (The TD has a representative of TD appear before the Senate Committee to oppose a piece of legislation having to do with bankruptcy cases for the reason that if that piece of legislation having to do with bankruptcy cases is adopted, X relieves AGEC of 77-B proceedings).

(Explan4) (The argument (G)-(A1') is a partial explanation of (A1'.1+R)). {from (Explan2),(C2),(I1) by (E4) and substitution}

(Explan5) (There is a partial explanation of (A1'.1.1)). {from (Explan3) by argument from existence}

(Explan6) (There is a partial explanation of (A1'.1+R)). {from (Explan3) by argument from existence}

It is important to note the following aspects of the above provisional specification of the Treasury's argument:

(i) It only shows that there are partial explanations for some of actions for which AGEC claims there are no satisfactory explanations; it does not show that these are satisfactory explanations. No criteria for the satisfactoriness of explanations in general and partial explanations in particular are involved; we have yet to discuss such criteria.

(ii) It only shows that there are partial explanations for some of the actions for which AGEC claims there are no satisfactory explanations. In particular, for reasons mentioned above, it does not show that there is an explanation for (a) the representative, as opposed to the Treasury, opposing the legislation on the ground (for the reason that) it would relieve AGEC of 77-B proceedings, or (b) for the representative or the Treasury opposing the legislation on the ground that it would be beneficial to AGEC.

(iii) The argument contains not only first-order propositions, but also many second-order propositions, such as the "Explan" and the "Inst" propositions.

(iv) All of the propositions in the argument are implicitly held to be beliefs of the Treasury's (this could be made explicit by prefacing them all by "The Treasury Department believes that"). It is essential that they be beliefs of the Treasury, i.e. reasons2, for rule (E4) to apply.

Let us continue with the construction of TD's argument. We have so far been concentrating on the argument corresponding principally to lines 40- 43, lines 151-157, and lines 329-330. We need now to go on and consider the lines 51-108,151-157,252-276,292-302, and 318-353.

In these sections, the TD expresses several goals of a higher level than the one we have studied so far (winning its suit), and puts forward a means-ends argument of a type which is much more complex than the types we have examined so far.

On an initial reading of these sections, we find the following new, not necessarily independent, TD super-goals mentioned (it will become apparent below why we call them "super-goals):

(Sup-G1) (TD wants (TD collects taxes due from AGEC)). (lines 57-58)

(Sup-G2) (TD wants (TD does TD's job fairly and honestly)). (lines 151- 153)

(Sup-G3) (TD wants (TD acts fairly towards AGEC)). (lines 62-64)

(Sup-G4) (TD wants (TD acts fairly towards the stockholders and creditors of AGEC)). (lines 62-64, 82-88)

(Sup-G5) (TD wants (TD collects the taxes due from AGEC with a minimum of hardship to creditors)). (lines 350-351)

We also find that the Treasury believes that it has several *methods* for collecting the taxes due from AGEC which it can use (lines 64-65):

(M1) (TD distrains the jeopardy assessment and seizes AGEC's property). (lines 65-67)

(M2) (TD forecloses the tax lien). (lines 72-74)

(M3) (TD wins TD's suit (77-B proceedings) against AGEC). (lines 74-76)

(M4) (TD makes a successful appeal to the Board of Tax Appeals). (lines 258-276)

Because of my lack of legal knowledge, I am uncertain about the inter- relationships between (M1)-(M4); in particular, I am not sure that (M2) and (M4) are entirely distinct. In what follows, however, I shall simply assume that they are all distinct. If it transpires that they are not distinct, it should be relatively clear how one would alter my analytical results to take account of my error.

The TD makes a comparative evaluation of the various methods (lines 64- 102, 255-276, 292-302) it has at its disposal to collect the tax (Sup-G1), by noting certain good- or bad-making characteristics of each and comparing their resultant values, and finally concludes that method (M3) is the best (most appropriate, most desirable) method (lines 81-82, 258- 263). Given that (M3) implies and explains (G) by virtue of the rationality-principle that an agent wants to adopt the best method available for attaining his goals, we begin to see the outline of the major explanatory argument put forward by the TD in the sections under study:

(1) Agent has super-goals Sup-G1...Sup-Gm.

(2) Agent can use methods M1...Mn to attain his super-goals.

(3) Method Mi is the best of the methods M1...Mn, because

(3.1) Method M1 has the set of good/bad-making characteristics C1.

.....

Method Mn has the set of good/bad-making characteristics Cn.

(3.2) By virtue of the set of evaluative rules or criteria EvC, method M1's characteristics, C1, warrant assigning it the ordinal value V1.

.....

By virtue of the set of evaluative rules or criteria EvC, method Mn's characteristics, Cn, warrant assigning it the ordinal value Vn.

(3.3) Vl > V1...Vk,Vm...Vn

(4) The agent wants to adopt method Ml; that is, the agent wants to perform those actions, to have those states obtain and those events occur which constitute the method Ml. (This is the agent's goal.)

(5) (1)-(4) constitute an explanation for (4), by virtue of the rule of explanation (E11):

(E11) An argument of the form {P1,P2,P3,P4}, where

(P1) states that an agent has goals G1...Gm

(P2) states that the agent can use (follow) methods M1...Mn to attain some or all of the goals G1...Gm

(P3) states that some method, Ml, a member of {M1...Mn}, is the best method of the set {M1...Mn}

(P4) states that the agent has as his (new sub-)goal the use of method Ml

is an explanation of (P4).

This rough outline raises at least two sets of questions: (a) what is the nature of "methods"? Can they be explicitated in terms of actions by the agent and causal sequences of actions and events? Is a method to attain a goal a "way to" or a "way of" attaining that goal (in terms of our distinction between means and action-instantiations)? And (b) what is more specifically involved in the process of comparative evaluation? What roles do the agent's (super-)goals play in determining value? What is the nature of the evaluative rules or criteria mentioned in the outline? We will discuss these questions in sequence.

(a) A method is a procedure or process for attaining an object (a goal). If an agent adopts a method to attain a goal, it is reasonable to suppose that (i) he performs at least one action, and (ii) that that action causes the goal (state, action, event, etc.) to obtain. A consequence of (ii) is that the action performed by the agent is distinct from the goal, and in cases where the goal is a general action by the agent, is not an instantiation of that general action (it is a "way to", not a "way of"). However, when an agent adopts a method, he may perform several actions in sequence, and these actions may be separated by causally linked chains of states or events which are not actions by the agent. Obviously, when we describe a sequence of actions, we must take time considerations into account. And when, e.g. in explaining, we describe the sequence of actions, states and events which occur when an agent adopts a method, we make use of implication relations as well as causal relations. For example, one method of getting a hard egg is to turn on the stove, put a pot of water on the hot stove, wait until the water boils, put the egg in the boiling water, and remove the egg after a certain amount of time. If we were describing what happens when an agent adopts this method, we might say something like:

(1) Agent turns on the stove at time t1. (Action)

(2) (Agent turns on the stove at time t1) -> (The stove is on at time t1). (Implication)

(3) (The stove is on at time t1) tcause (The stove is hot at time t2). (Causal statement)

(4) (The stove is hot at time t2). (from (2),(3))

(5) Agent puts a pot of water on the stove at time t2. (Action)

.....and so on.....

So a method is a *course of action* open to the agent: an ordered set of actions, {A1 at t1 (, A2 at t2, ..., An at tn)}. The times in question need not, of course, be punctual; they may be durations, which may overlap, but with the restriction that if action An continues from tq to tr, and action An+1 continutes from ts to tt, tq be earlier than ts. What happens when an agent adopts a method, a course of action, and what a deliberating (planning) agent believes will happen when he adopts a course of action, is a sequence of causally linked actions, states and events, an "*ASE-sequence*". When ASE-sequences are described in explanations, they are described by sequences of propositions which include not only causal statements but also inferential statements.

(b) The comparative evaluation by an agent of different methods or means to attain a goal or a set of goals requires the assignment of ordinal values to the different methods or means (in the next few paragraphs we shall use the terms "method" and "means" interchangeably). The value attributed to a means depends on at least three distinguishable factors: (i) the agent's *preferences between his various goals* (if he has more than one goal), expressed as ordinal values *, (ii) the agent's perception of the mean's *comparative efficiency, ie. the comparative probability that it will produce the desired goal(s)* , and (iii) its *intrinsic desirability for the agent, i.e. his preference for the ASE- sequence corresponding to that method without consideration of the sequence's outcome or its efficiency in producing that outcome* (this is again expressed as an ordinal value). * We will discuss each of these in

---

* I assume that utilities can only be measured ordinally, based on my knowledge of of economic theory, which points out that utilities can only be measured on an interval or ratio scale in certain very restricted conditions, if at all. Now it may be that when people are deliberating about value and utilities, they make use of ratio scales, though, upon introspection, I find that hard to believe. If that is the case, of course, many of the indeterminacies mentioned below would no longer be a problem.

---

* The following discussion is largely inspired by the excellent discussion of behavioral notions of value within a system-theoretic framework in [Russell Ackoff and F.E.Emery, "On Purposeful Systems", Tavistock, London, 1972, Chapter III].

turn.

(b.i) The agent is considering various courses of action open to him. These courses of action have different outcomes; initially one might suppose that these outcomes are the attainment of one or more of the agent's goals (we assume that the agent does not consider a course of action if it does not lead to the fulfillment of at least one of his goals). Other things being equal, an agent will assign a higher value to a course of action which leads to the attainment of G1 than to a course of action which leads to the attainment of G2, if he prefers G1 to G2.

It is important to note that, strictly speaking, there are many more courses of action open to an agent than there are means at his disposal, if one uses the term "means" in a natural way. Suppose that I have two means at my disposal to attain a goal state GS (being President): M1 - large-scale bribery, and M2 - campaigning. The courses of action open to me are, strictly speaking, all possible sequences of means: M1 only, M2 only, M1 and M2 simultaneously, M1 followed by M2, M2 followed by M1, M1 followed by M1 and M2, and so on ad infinitum. There may be causal inhibiting or enabling relationships between these means, of course, which would decrease the possible courses of action, but in most cases, the number of courses of action is far greater than the number of means. In what follows, we will - and I believe this is a reasonable step, ignore the use of multiple means and the sequencing of means, and simply say that if an agent has X number of means, then he has X number of courses of action open to him; in other words, we will establish a one-to-one correspondence between means and courses of action.

(A short terminological clarification is in order here. There is a looseness in our talk about goals which has not been troublesome up to this point, but which should be pointed out now. Suppose an agent has a goal. What this amounts to is that the agent wants some action to be performed, some state to obtain, some event to occur, or some combination of these. So our description of a state of affairs in which an agent has a goal is a statement of the form "AGENT wants (A/S/E)". Strictly speaking, a goal is an A/S/E such that some agent(s) wants it. And when we speak about the adoption of a method leading to the attainment or fulfillment of an agent's goal, we observe this strict usage, for what we have in mind is something like "(Method tcause A/S/E1), (Agent wants A/S/E2), and (A/S/E1 = A/S/E2)". But we do not always observe this strict usage. In particular: (a) Sometimes when we talk about "the agent's goal", we refer by that phrase not to the A/S/E such that the agent wants it, but to the state of affairs in which the agent wants A/S/E. And (b) sometimes we say things like "John is aware of Mary's goal", meaning thereby that John is aware that Mary wants X. To repeat: sometimes the word "goal" is used strictly to refer to an A/S/E such that some agent(s) wants it, but sometimes it is used loosely to refer to a state of affairs in which some agent wants some A/S/E. This ambiguity is often useful, and resolvable in context without difficulty. We shall play upon it when needed below, hopefully without causing any confusion. The possibility of confusion should be lessened by this explicit warning, however.)

This suggests that if methods M1...Mn produce, respectively, attainment of goals G1...Gn, and the goals have the ordinal values n...1 (where n is highest and 1 lowest), then the methods will be assigned the "preliminary type-1" ordinal values n...1.

This suggestion is too simple as it stands, however, for two reasons. (1) As there are

multiple goals, a particular course of action or method may lead to the attainment of more than one goal. And (2) a method which leads to the attainment of one goal may lead to the non-attainment of another, and this fact must be taken into account when assigning it a value.

(1) What if a course of action leads to the attainment of more than one goal? We cannot simply assign it a preliminary type-1 value which is the sum of the values assigned to the two goals in question, for ordinal measurements cannot always be usefully summed in this context (i.e. if A>B>C, it does not follow that (B + C) is larger, smaller or equal to A, though it does follow that (B + C) > B).

Consideration of this difficulty shows that we must distinguish between the outcomes of the courses of action open to the agent and the agent's goals. This distinction is clear on the face of it, of course: statements of goals are of the form "A wants X", whereas statements of outcomes are of the form "X". The courses of action open to the agent have various outcomes, and these outcomes (1) may only partially fulfill the agent's goal(s) *, and (2) may to some degree fulfill one or more of the agent's goals. To illustrate (1), suppose that an agent wants to acquire a house; he may have a course of action open to him which would lead to his acquiring a delapidated shack, and we would say that that outcome would represent only a partial fulfillment of his goal. Judgments about the degree to which an outcome is the attainment of a goal are very complex, as they involve judgments about the similarity or closeness of states of affairs along relevant or important dimensions.

It is clear that the preferential ranking of outcomes is a very complicated function of the preferential ranking of goals and the relationships between outcomes and goals, or n-tuples of goals. For the sake of facilitating our analytical work, we will make the following simplifying assumptions, which will not invalidate our analysis of our particular text:

- Outcomes and goals are identical.

- The goals used in the analysis will be so defined as to comprise some of the possible combinations of the goals directly derived from the text. For example, suppose that the text provided direct evidence for the TD's having two goals: (G1) collecting the tax money, and (G2) being fair In our analytical representation of the TD's means-selection argument, we would postulate the following three goals: (G1) collecting the tax money; (G2) being fair, and (G3) collecting the tax money and being fair. However, in one's representation of an explanatory argument, one does not, for

---

* To put this more precisely: an agent has a goal - (Agent wants A/S/E1) - and a method leads to an outcome - (Method tcause A/S/E2). The outcome may be more or less "close" to the goal - A/S/E1 may be more or less "close" to A/S/E2. The only area of the philosophical literature which I am aware of and which (summarily) discusses the degree of closeness of actions, events or states of affairs is David Lewis's discussion of the formal semantics of counterfactual conditionals and modal logic, in which he makes use of a "proximity" relation between possible worlds.

reasons of economy, want to represent all of the possible combinations of goals. Rather, one only represents the textually specified goals and those combinations of them which play an actual role in the explanation.

In specifying the rankings of goals, we will make use of any applicable rules of ordinal calculation, such as the rule that (A • B) is preferable to either A or B alone.

(2) Given that there are many goals, a given means may lead to the attainment of some goal(s) and to the non-attainment of others. Suppose we have two goal-states, GS1 and GS2, and a means, M, causes GS1 but not GS2. It is not sufficient to say that the preliminary type-1 value of M is dependent on the value of GS1; for it is in fact also dependent on the value of the state ~GS2.

This seems to me to be intuitively obvious. However, taking the values of non-attainment of goals into account in addition to the attainment of goals makes surprisingly little difference in the calculations which allow us to calculate the ordinal type-1 values of means. Consider the following two cases:

	GS1	~GS2	GS2	~GS2
V:	3	1	4	2
M1	3		4	
M2	3			2
M3		1	4	
M4		1		2

	GS1	~GS1	GS2	~GS2
V:	3	2	4	1
M1	3		4	
M2	3			1
M3		2	4	
M4		2		1

These two tables show the values of the attainment and non-attainment of two goals (4 is best and 1 is worst). M1 is a means that leads to the attainment of GS1 and GS2, M2 leads to GS1 but not GS2, M3 leads to GS2 but not GS1, and M4 leads to neither GS1 nor GS2. In these two cases, and in all other assignments of values to goal-states which obey the plausible constraint that the value of a goal obtaining must be greater than the value of a goal not obtaining, the only difference that considering the values of non-attainment in addition to the values of attainment makes is that if less than the full set of means is considered, then the ranking of them is incomplete. Thus in the first case, if one considers only M1-M3, then if one uses only the values of attainment in one's calculations, one obtains M1>M3>M2, whereas if one uses both the values of attainment and the values of non-attainment, one obtains only M1>M2,M3 (it is indeterminate whether M2>M3, M2<M3, or M2=M3). If one considers all the means M1-M4, then one gets a complete ranking of them, whether one takes account only of the values of attainment, or whether one takes account of both the values of attainment or of non-attainment.

Given this fact, and the fact that it makes no difference to the analysis of our particular means-selection argument whether we bring in non-attainment values or not, we will not pursue this matter furthur here.*

(b.ii) If two means lead to the same outcome, a rational agent will, ceteris paribus, prefer that means which he perceives as having the highest probability of producing

the goal. This is the simplest case. In the more general case, if means M1 leads to G1 with probability p1, and means M2 leads to G2 with probability p2, then a rational agent will prefer the means that has the highest expected instrumental value; and this expected instrumental value of the means is the product of p, the probability that it will produce the goal, and v, the value of the goal. So means are assigned "preliminary type-2" ordinal values which are their expected instrumental values, and these expected instrumental values are the products of the preliminary type-1 ordinal values of the goals the means lead to and the subjective probabilities of success of the means.

There is a clear problem involved in this derivation of expected instrumental values: the values of the goals are ordinal, and the probabilities are either arranged on a ratio scale, or else (given that they are subjective probability estimates) are themselves ordinal. So we can only draw conclusions about comparative instrumental values in certain limited cases. For example, if p1 >/= p2, and v1 > v2, then we can conclude that (p1 x v1) > (p2 x v2); however, if p1 > p2, but v1 < v2, we can draw no conclusions about the equality or inequality of (p1 x v1) and (p2 x v2). So instrumenal values can often be indeterminate.

(b.iii) Means are also assigned a set of preliminary type-3 ordinal values which represent the intrinsic desirability or worth, in the eyes of the agent, of the ASE-sequences which correspond to them. Suppose that I have two methods for communicating a message to someone who lives on top of a mountain: climbing the mountain and sending smoke signals. If I am making a comparative evaluation of these methods, one set of considerations which I will have to take into account is the intrinsic attractiveness of the two methods per se. Climbing the mountain, I might think, would be dangerous and laborious, whereas sending smoke signals would be much less laborious, novel, and would allow me to learn something about a communicative media which I have always been curious about. Not only the actions comprised in the ASE-sequence corresponding to a method are evaluated, but also the states or events. These actions, states or events have certain relevant properties, and they are evaluated by applying rules ("All actions with properties P, Q, R... are right/wrong.") and evaluative standards ("All states or events with properties P, Q, R... are good/bad). * Assigning an intrinsic value to an ASE-sequence as a whole is in itself a complex evaluative task, as it involves evaluating each component in the light of its relevant properties and various rules or criteria (which interact in various complicated ways **) and generating a resultant value. We shall not attempt to specify the rules and

---

*There are many fascinating questions about the value of the states which are the complements of goal states. Can one make sense of a notion like "all states in which GS does not obtain", intuitively speaking? Is there any kind of correlation between the magnitude of the value of a GS and of ~GS? Does disappointment enter in?

---

*Both rules and evaluative criteria are mentioned here. I will not go into the long-standing controversy about the distinctions between rules and criteria, deontological theories of obligation and theories of value, and between the right and the good. Cf. [Sir David Ross, "The Right and the Good", Clarendon Press, Oxford, 1930], and most introductory moral philosophy textbooks.

procedures used to assign such values; to do so would require discussing large areas of meta-ethics.

It might well be objected that many of the rules and criteria used to evaluate the intrinsic desirability of ASE-sequences could equally well be represented as goals, in which case they would influence the preliminary type-1 and type-2 values of the methods, and not their type-3 values. Suppose that I have a goal G, and two methods for attaining it, M1 and M2. Suppose that I feel M2 is less intrinsically desirable than M1, because M2 would involve breaking the law, whereas M1 would not, and I adhere to a rule of conduct or norm, R, which is "Do not act illegally.". Why could I not simply say that I have two goals - G and obeying the law (G') - and M1 has a higher type-1 value because it leads to (G + G'), whereas M2 only leads to (G)? I can only reply that there is no unassailable reason why one should not do this. However, when people are delibe[illegible] about particular practical decisions, they tend to regard as goals those goa[illegible]eirs which are more particular to their particular decision situation, and reg[illegible] as rules or criteria of intrinsic desirability of methods those general goals of thei[illegible]hich are involved in a much wider range of decision situations (such as staying alive, acting virtuously or legally). It is the generality of rules and goals which is of importance here.

The final values assigned to the methods (i.e. the final output of the comparative evaluation process) are a function of their preliminary type-2 values and their preliminary type-3 values. As mentioned above the preliminary type-2 values are themselves a function of the preliminary type-1 values and the subjective probabilities of their success in producing the goals/outcomes which they might lead to.

In light of the above discussion, let us expand our outline of of a means-selection argument.

(1) Agent has (super-)goals Sup-G1...Sup-Gn:

(Sup-G1) (Agent wants A/S/E1)

.....

(Sup-Gm) (Agent wants A/S/Em)

**For example, rules of obligation are of different types and levels (cf. [Joseph Raz, "Practical Reason and Norms", Hutchinson and Co., London, 1975]) and may "over-ride" each other (cf.[Roderick Chisholm, "Practical Reason and the Logic of Requirement", in Stephan Korner, ed., "Practical Reason", Yale University Press, New Haven, 1974]).

and combinations of goals:

(Sup-Gm+1) (Agent wants (A/S/E1 + A/S/E2))

.....

(Sup-Gn) (Agent wants (A/S/Ex + ...)

(2) The agent has certain methods, M1...Mp, at his disposal:

(M1) It is possible that M1 (where M1 is an action by the agent, or a sequence of actions by the agent).

.....

(Mp) It is possible that Mp

(3) These methods would lead to the attainment of certain goals, or combination of goals:

(MC1) M1 tcause A/S/Ex

.....

(MCp) Mp tcause A/S/Ez

(4) The goals which the methods would attain have certain ordinal values:

(GV) V(A/S/Ex) >,=,< ..... >,=,< V(A/S/Ez)

(5) The means have certain probabilities of leading to the attainment of the goals they might tcause:

(MCP1) M1 tcause A/S/Ex with probability p1

.....

(MCPp) Mp tcause A/S/Ez with probability pp

(PR) p1 >,=,< ..... >,=,< pp

Note: if A/S/Ex is a combination of members of A/S/E1...A/S/Em, then the probability of a method producing that combination is a product of the probabilities of the method producing each member of the combination, assuming independence.

(6) The preliminary type-2 values of the means are the products of the probabilities with which they will lead to the goal actions/states/events they might lead to, and the values of those goal actions/states/events. They are ranked aaccordingly (by virtue of the rules of ordinal calculation):

(MV2) V2(M1) = (p1 x A/S/Ex) >,=,< ..... >,=,< V2(Mp) = (pp x A/S/Ez)

(7) The methods M1...Mp, or the ASE-sequences which correspond to them, have certain attributes which are relevant to their intrinsic desirability or value.

(8) By virtue of evaluative rules, and the attributes mentioned in (7), the means are assigned preliminary type-3 values:

(MV3) V3(M1) >,=,< ..... >,=,< V3(Mp)

(9) By virtue of a furthur set of evaluative rules, and the principles of ordinal calculation, a the final comparative values of the means is determined from (MV2) and (MV3):

(FMV) FV(M1) >,=,< ..... FV(Mp)

(10) The agent wants to perfom the actions corresponding to the method which has the highest final value. (This is the agent's new goal.)

(11) (1)-(10) constitute an explanation of (10) by virtue of (E5).

Let us now go over the dialogue and consider what the TD says about each one of the four methods open to it, attempting to categorize these remarks within the above framework.

(M1) - distraining the jeopardy assessment and seizing AGEC's property.

a) The Treasury has attempted to avoid using M1 if possible (lines 67-68). This simply tells us that after comparative evaluation, the TD has assigned a low ordinal value to M1; furthur remarks are must be examined before we can surmise why.

b) M1 is an extreme method. We deduce this because both M2 and M3 are termed more moderate (lines 72-76). I believe that this extremism has a bearing on its intrinsic desirability. It could be argued, of course, that to be extreme is to be unfair, so that the extremism of M1 means that it does not lead to (Sup-G2)-(Sup-G4), and consequently has a lower type-1 and perhaps type-2 value. But extremism can cover many other features besides unfairness, and in doubt, I prefer to say that it influences M1's type-3 value.

c) M1 would not give the creditors and stockholders of AGEC a look-in on the proceedings, and would not give the creditors a chance to be heard. If the creditors do not have a chance to be heard, they might be left out in the cold. (lines 81-88, 292-295) It seems reasonable to assume that if the TD causes the creditors and stockholders of AGEC to have a look-in on the proceedings, that would be an instance of the TD's acting fairly towards the creditors and stockholders (Sup-G4). It also seems reasonable to say that if the TD does something which might result in the creditors being left out in the cold, then it is not attaining (Sup-G5), which

involves imposing a minimum of hardship of the creditors. So M1 has a lower type-1 value in so far as it does not lead to (Sup-G4), and has a lower type-2 value in so far as it has little chance of attaining (Sup- G5).

d) M1 may be used while the 77-B proceedings are going on (lines 254-257). This remark is concerned with causal enabling and inhibiting relationships between methods. If we are to take such relationships into account, we must complicate our framework, which supposes independent and distinct methods. As an example of such a modification, suppose that we had methods M1, M2 and M3, and that M1 inhibited M2 and M3. In that case we would have to say that we had the following methods: (M'1) - M1, (M'2) - M2, (M'3) - M3, and (M'4) - M2 • M3.

(M2) - foreclosing the tax lien.

a) M2 is moderate, more moderate than M1 (lines 72-74). This means that it is, in this respect at least, intrinsically more desirable than M1.

b) M2, like M1, would not give the creditors and stockholders of AGEC a look-in on the proceedings, a hearing, and might result in the creditors being left out in the cold (lines 81-88, 292-295). So, like M1, M2 has a lower type-1 value in so far as it does not lead to (Sup-G4), and a lower type-2 value in so far as it has little chance of attaining (Sup-G5).

c) M2, unlike M1, cannot be used while the 77-B proceedings are going on (lines 254-257).

d) M2 is a method which the TD had already started to use prior to the time of the dialogue (lines 72-73). (Presumably the TD is no longer using it at the time of the dialogue, because all of its energies are being absorbed by the 77-B proceedings ((c) above). It is reasonable to suppose that since the TD adopted method M2 but then abandoned it in favour of M3 (77-B suit), but did not adopt either M1 or M4, the final evaluation of M1-M4 gives the following preferential ranking of methods: M3 > M2 > M1 >/= M4. This inference depends on (a) the principle of rationality according to which a rational agent adopts the method he deems best, and (b) the assumption that at the time the TD filed a bill to foreclose the tax lien it was not aware of or capable of acting upon the possibility of successfully waging a 77-b suit.

(M3) - the 77-B suit against AGEC, if successful.

a) M3 is the best method for determining the tax liability and collecting the tax (lines 81-82, 258-259, 261-263, 320-322). This is evidenced by the fact that it is the method which the TD has in fact adopted, assuming again, as one does in reason2-explanations, that they are rational. So the TD's comparative evaluation of methods assigns the highest final ordinal value to M3.

b) M3 is moderate, more moderate than M1 (lines 74-76). This means that in this respect it is intrinsically more desirable than M1.

c, M3 gives the creditors and stockholders of AGEC a look-in on the proceedings (lines 82-88, 292-295), and is therefore fair to them; it has a higher type-1 value than M1, M2, and M4, in that it leads to (Sup-G4). It also will probably not leave the creditors out in the cold, so it has a higher type-2 value than M1, M2, and M4 with respect to attaining (Sup- G5).

d) M3 inhibits M1, M2, and M4 (lines 256-257).

e) M3 allows the tax liability to be determined more quickly than M4 (the use of the Board of Tax Appeals) (lines 25[illegible]). Since determining the tax liability is a pre-requisite for collecting the tax money, M3 will allow a quicker collection of the tax money [illegible] might think that this greater speed makes M3 preferable to M4 [illegible] because perhaps collecting a given amount of money sooner than later means that one collects more money in real terms (discounting). However, this is not the salient reason why speed is of the essence. Consider lines 274-276: "Section 77- B has the effect of preserving the assets. By the time the Board of Tax Appeals gets it, there may be nothing left to collect.". The assets are disappearing. If the TD attempts to collect the tax money due some time in the future, they may not succeed. So M3 is preferable to M4 in that it is much more likely to lead to effective collection of the tax money due; M3's expected instrumental value with respect to (Sup-G1) is higher than M4's.

f) M3 allows the TD to obtain from AGEC the information that is necessary to compute the tax and then collect it (lines 97-108). 77-B involves a trusteeship, which would allow easy gathering of the information; other methods do not, and they would involve a lengthy legal coercion of the uncooperative AGEC to give up the information, and meanwhile the assets would be disappearing (cf. (e) above). So this lack of cooperation by AGEC in the matter of providing information is another reason why M3 has a higher probability of success in attaining (Sup-G1) than do M1, M2, and M4.

g) M3 enables the TD to collect less than the full amount of tax (lines 88- 90), whereas M1, M2, and M4 do not. This at first seems to count against it; M3's outcome does not involve the complete attainment of (Sup- G1). Certainly AGEC perceives this reason as counting against it (lines 92-95). However, consider lines 350-353. It seems that if the TD collects the full amount of the tax, it may be imposing more hardship on creditors (who would have little left to collect), than if it collects less than the full amount due. So it seems that (Sup-G1) might conflict with (Sup-G5). So far, we have not considered incompatibilities between goals/outcomes. Rather than add the complications that such considerations might bring, let us rather add a new goal, (Sup-G6), (TD wants (TD can accept less than the full amount of the tax money due)), and say that M3, but not M1, M2, and M4 lead to it. And to avoid the bizarreness of saying that M3 leads both to the TD's collecting (the full amount of) the tax and being able to collect less than the full amount of the tax, we should modify (Sup-G1) to read (TD wants (TD can collect the full amount of the tax money due from AGEC)).

(M4) - making a successful appeal to the Board of Tax Appeals.

a) M4 would not give the creditors and stockholders of AGEC a look-in on the proceedings, would not give the creditors a chance to be heard, and might leave the creditors out in the cold (lines 81-88, 292-295). So, like M1 and M2, it would not lead to (Sup-G4) and has a smaller chance of leading to (Sup-G5).

b) M4 is inhibited by M3, as is M2 (lines 255-257).

c) M4, unlike M3, does not have the effect of preserving the assets (line 274), does not allow the Court to determine the tax liability quickly (lines 258-261), and therefore may not allow the TD to collect tax money due. M4 is less probable than M3 to lead to (Sup-G1).

So it is possible and, hopefully, useful, to categorize the remarks made by the TD about the methods open to it using the concepts drawn from our outline of comparative evaluation procedure. However, these remarks are far from giving us all of the information required by our model. Some evaluatively relevant attributes are predicated of some methods but neither predicated nor pronounced lacking in others. For example, we know that M1 is more extreme than M2 and M3, but nothing is said about the extremism or moderation of M4. Probabilities are not specified with any degree of precision, either: we are told that M1, say, could leave the creditors out in the cold, but this only tells us that the probability of M1's leaving the creditors out in the cold, and consequently failing to impose only a minimum of hardship on them, is greater than 0 but less than 1. We shall see that these indeterminacies do not, nevertheless, result in AGEC's means-selection argument being unsound, or failing to justify the desired conclusion.

Below, we will display our construction of AGEC's explanatory means-selection argument. First we will list some important assumptions and notational conventions used in its formulation.

We have made the following assumptions:

1) If the text does not justify one's concluding that a given mean or method does or does not lead to (tcause) attainment of a goal(s), we assume that it does not.

   We will also not consider the value of non-attainment of goals, for the reasons mentioned above.

2) If the text simply asserts or implies that a mean tcauses attainment of a goal(s) without specifying any probability, we assume that it tcauses the goal action/state/event with an default "high" probability. Usually, when we make a causal statement, we don't think that the probability that the cause produces the effect 1, but we do think that it is reasonably high, certainly greater than .5. There is almost a conversational postulate to the effect that if we believe the probability to be less than .5 (or perhaps less than the probability with which other causes would produce the same effect), we should mention that fact.

3) If the text provides reasons why a means might not succeed in producing a goal(s), we simply say that the probability of the means tcausing the goal is "low", i.e. less than the default "high" probability.

4) If a given means, M, tcauses G1 with probability p1, G2 with probability p2, ..., Gn with probability pn, we assume that theses occurrences are statistically independent, i.e. that p(Gi/M) = p(Gi/M,Gx) for all x.

5) We will assume, as we stated above, that the courses of action open to the TD are identical with the means at its disposal: we will not consider combinations of means to be means.

6) Though we do mention that one means (77-B suit) inhibits the other means, we do not, because of what was said directly above in (5), make use of this fact in our comparative evaluation calculations.

These are of course extremely simplifying assumptions, which would render our analytical technique incapable of capturing the subtleties of many means-selection arguments. They do allow us to handle the argument we are concerned with here, however.

In the interests of expository brevity, we have adopted the following notational conventions:

1) Instead of writing out several statements which only vary with respect to one of their components - such as statements of the form AX, BX, CX - we write a single statement comprising commaed lists - such as A,B,C X

2) When a goal is first mentioned, it is specified by a statement of the form: (Gn) Agent wants A/S/E. Thereafter, however, we will simply use the label to stand in for the A/S/E in question.

3) A similar abbreviation is used in the case of means. When a means is first mentioned, it is specified by a statement of the form: (Mn) it is possible that (Agent-Action). Thereafter, however, its label will stand in for the Agent-Action in question.

In addition, we use comma-ed lists in rankings to signify that the values referred to by the expression joined by commas are unordered. Thus A > B,C > D means that A is larger than both B and C, that B and C are both larger than D, but that it is not determined whether B is larger than, smaller than, or equal to C.

We will make use of the following new rules:

RG1: If (Agent wants A/S/E1) ... (Agent wants A/S/En), then (Agent wants (A/S/E1 and ... A/S/En).

R3: If X tcause A, and A is an instance of GA, then X tcause GA, where A is a specific action, and GA is the same action described in a more general manner

RVG1: If (Agent wants A/S/E1) and (Agent wants A/S/E2), then - in the eyes of the agent - V(A/S/E1 • A/S/E2) > V(A/S/E1),V(A/S/E2).

Lastly, we will omit the attribution of type-1 preliminary values, and go on immediately to the attribution of type-2 values.

Here is our construction of AGEC's explanatory means-selection argument:

(Sup-G1) (TD wants (TD collects the full amount of tax money due from AGEC)). (lines 57-58)

(Sup-G2) (TD wants (TD can collect less than the full amount of tax money due from AGEC)). (lines 88-90)

(Sup-G3) (TD wants (TD acts fairly towards the stockholders and creditors of AGEC)). (lines 62-64, 82-88)

(Sup-G4) (TD wants (TD imposes a minimum of hardship on the creditors of AGEC)). (lines 350-351)

(Sup-G1•4) (TD wants ((TD can collect the full amount of tax money due from AGEC) and (TD imposes a minimum of hardship on the creditors of AGEC))). {from (Sup-G1),(Sup-G4) by RG1}

(Sup-G1•2•3•4) (TD wants ((TD collects the full amount of tax money due from AGEC) and (TD can accept less than the full amount of tax money due from AGEC) and (TD acts fairly towards the stockholders and creditors of AGEC) and (TD imposes a minimum of hardship on the creditors of AGEC))). {from (Sup-G1),(Sup-G2),(Sup-G3),(Sup-G4) by RG1}

(M1) It is possible that (TD distrains the jeopardy assessment and seizes AGEC's property). (lines 65-67)

(M2) It is possible that (TD forecloses the tax lien). (lines 72-74)

(M3) It is possible that (TD wins TD's 77-B suit against AGEC). (lines 74-76)

(M4) It is possible that (TD successfully appeals to the Board of Tax Appeals). (lines 258-276)

(MC1) (M1),(M2),(M3),(M4) tcause (Sup-G1). (lines 65-67, 72-76, 258-261)

(MC2) (M1),(M2),(M3),(M4) tcause (Sup-G4). Because (lines 82-88):

- (M3) tcause (the creditors of AGEC have a chance to be heard).

- (the creditors of AGEC have a chance to be heard) tcause ~(the creditors of AGEC are left out in the cold). (with high probability)

Therefore, (M3) tcause ~(the creditors of AGEC are left out in the cold). {by transitivity of tcause}

- If ((M3) tcause ~(the creditors of AGEC are left out in the cold)) then(pf) ((M3) tcause (AGEC imposes a minimum of hardship on the creditors of AGEC)).

Therefore, (M3) tcause (Sup-G4). (with high probability)

- (M1),(M2),(M4) tcause ~(the creditors of AGEC have a chance to be heard).

- ~(The creditors of AGEC have a chance to be heard) tcause ~(the creditors of AGEC are left out in the cold). (with low probability, but some probability nevertheless)

Therefore, (M1),(M2),(M4) tcause ~(the creditors of AGEC are left out in the cold). {transitivity of tcause}

If ((M1),(M2),(M4) tcause ~(the creditors of AGEC are left out in the cold)) then(pf) ((M1),(M2),(M4) tcause (AGEC imposes a minimum of hardship on the creditors of AGEC)).

Therefore (M1),(M2),(M4) tcause (Sup-G4). (with low probability)

(MC3) (M3) tcause (Sup-G2). (lines 88-90)

(MC4) ~((M1),(M2),(M4) tcause (Sup-G2)). (lines 88-90)

(MC5) (M3) tcause (Sup-G3). Because (lines 82-85):

- (M3) tcause (the creditors and stockholders of AGEC are given a look-in on the proceedings by TD).

- (The creditors and stockholders of AGEC are given a look-in on the proceedings by TD) is an instance of (TD acts fairly towards the creditors and stockholders of AGEC).

Therefore, (M3) tcause (Sup-G3). {by R3, and a (unspecified!) theory of fairness or justice}

(MC6) (M1),(M2),(M4) tcause (Sup-G1•4). {from (MC1)-(MC5)}

(MC7) (M3) tcause (Sup-G1•2•3•4). {from (MC1)-(MC5)}

(GV)

V(Sup-G1•2•3•4)

> V(Sup-G1•4)

> V(Sup-G1),V(Sup-G2),V(Sup-G3),V(Sup-G4). {from (Sup-G1...1+2+3+4) by (RVG1)}

(MCP1) (M1) tcause (Sup-G1) with high probability ph, (Sup-G4) with low probability pl. (lines 65-67, 82-88)

(MCP2) (M2) tcause (Sup-G1) with high probability ph, (Sup-G4) with low probability pl. (lines 72-74, 82-88)

(MCP3) (M3) tcause (Sup-G1) with high probability ph, (Sup-G2) with high probability ph, (Sup-G3) with high probability ph, (Sup-G4) with high probability ph. (lines 74-76, 88-90, 82-85, 86-88, 350-351)

(MCP4) (M4) tcause (Sup-G1) with low probability pl, (Sup-G4) with low probability pl. (lines 258-261, 274-276, 97-108, 86-88)

(PR) ph > pl

(MV2)

V2(M3) = (ph x V1(Sup-G1)) + (ph x V(Sup-G2)) + (ph x V(Sup-G3)) + (ph x V(Sup-G4))

>

V2(M1) = V2(M2) = (ph x V(Sup-G1)) + (pl x V(Sup-G4)

>

V2(M4) = (pl x V(Sup-G1)) + (pl x V(Sup-G4)). {by ordinal calculation}

(DA) (M1) is extreme; (M2),(M3) are less extreme. (lines 65-76)

(V3R) For all x,y, if y is less extreme than x, then(pf) V3(x) < V3(y).

(MV3) V3(M2),V3(M3) > V3(M1)

(MFV) FV(M3) > FV(M2) > FV(M1),FV(M4). {from (MV2),(MV3), and (RVM1)}

(G) TD wants (M3) - i.e. TD wants (TD wins TD's 77-B suit against AGEC).

(Explan) (Sup-G1)-(G) is an explanation of (G). {by (E5)}

The above means-selection argument is satisfying in that it yields a final ranking of means in accordance to that which we deduced from the text (cf. (M1-(a)), (M2-(b)), and (M3-(a)) above).

The argument (Sup-G1)-(Explan), appended to the argument (G)-(Explan6), forms our complete representation of the TD's means-ends argumentation.

(iv) General Comments About the Argument. The argument which we have been studying is a lengthy practical argument displayed as a part of an argument to the effect that there is an explanation (i.e. an explanatory argument) for an agent performing certain actions. In constructing both arguments, we have postulated certain plausible rules and argument forms which were useful for this particular analysis. We have not had the time to carefully consider the general validity of these rules, to systematically search for counter-examples for each, and progressively modify them in light of the counter-examples until they could reasonably be claimed to be valid. Such a task would require many furthur studies.

We will at present simply list the general problems which the present analysis has touched upon, some of which have been discussed in the literature to some extent, and briefly discuss the relationships between our analysis and those problems.

(I) The nature of *explanation*.*

(1) How does the form and content of explanations, of that which explains (the "explanans"), vary with respect to the form and content of that which they explain (the "explanandum")?

(2) What is the purpose of explanation? When do people feel called upon to seek for or generate explanations, particularly in dialogues? How does the form and content of explanations vary with respect to the purposes they serve, the functions they fulfill?

(3) What are the criteria for satisfactoriness of explanations? How do these criteria vary with (a) the form and content of the explanation, (b) the purposes with which the explanation is sought or generated?

With respect to (1), the explanation which we have been considering is certainly determined by the fact that it's explanans is a purposeful action taken by a presumably rational, deliberating agent. If the explanation had had as it's explanans an event in the physical world, it would have been very different both in form and content.

With respect to (2), the primary purpose of the explanation we have examined was to render an agent's action intelligible; its secondary purpose was perhaps to justify that action by showing that the reasons for it were socially condoned. Clearly, if the purpose of an explanation is prediction and control, as would be the case of an explanation put forward by an engineer, its form and content would be influenced by that purpose, and would be different from the form and content of the explanation we studied. A purely justificatory explanation would again be different. Even an explanation primarily aimed at rendering an agent's action intelligible may be shaped by furthur subsidiary purposes, such as wanting to sympathize with and help the agent, or as wanting to formulate worthy goals and courses of action for oneself (not the agent).

---

*For an introduction to the literature in this area see Stephan Korner, ed, "Explanation", Yale University Press, New Haven, 1975].

With respect to (3), though we have mentioned several times that the TD is called upon by AGEC to provide a satisfactory explanation, our argument does not show the explanation provided to be a satisfactory one. The reason that we have not been able to show that it is such is that we lack criteria of satisfactoriness for explanations of purposeful action. Intuitively, it seems as though the TD's and AGEC's criteria vary: AGEC wants a justificatory as well as an explanatory argument for the action of the TD, whereas the TD seems satisfied by what is for the most part a purely explanatory argument.

(II) The nature of *practical argumentation.*

Practical argumentation is argumentation which links up (a) goals, states, events, and actions (which may be actions of the agent) desired by an agent, and (b)

- other (sub-)goals of the agent, or

- actions of the agent,

through a system of beliefs, motives, intentions, values (criteria, standards), rules (rules of reasoning, of inference, of comparative evaluation, of verification, rules of thumb, rules of obligation, norms), and so on. Practical arguments make use of logical, instantiation, and causal relationships, as well as a set of rules about what values, beliefs and inferences one can impute to another person.

As is clear from our analysis, a study of practical argumentation evokes a host of the most intricate logical and philosophical difficulties, most of which we have brutally ignored in the interests of expediency. Our ignorance is not blissful, however. We will list the following areas of study which must be considered by anyone seriously attempting to study practical argumentation:

- the logic of statements in which we impute beliefs to others, which differs ordinary logic; two ways in which it differs is that it has to deal with the referential opacity of belief contexts and in that it has to place a limit on the length of chains of inferences (if a person believes A, B...Z, B -> C, C -> D, ... , Y -> Z, he does not necessarily believe A -> Z, because he may not have run through that long an inference); ***

- the nature of value, the different kinds of value and their inter- relationships, evaluation processes, the relationship of behavior to action, and the logic of statements in which we attribute values to others;

- the semantics and logic of our talk about actions: What is an action? How is it different from a state or event? What are the semantic and logical differences between (the description of) a general action and (the description of) a specific action? What are our identity criteria for actions? What does it mean to say that one action is an instantiation of another?

***One lead-in to the literature in this area is: [Nicholas Rescher, "Epistemic Modality: The Problem of A Logical Theory of Belief Statements", in Nicholas Rescher, ed, "Topics in Philosophical Logic", D.Reidel, Dordrecht-Holland, 1968].

- the semantics and logic of our talk about states and events (questions similar to those directly above);

- the nature of our intuitive notion of causality, and the logic of our statements about causality;

- the nature and rules of the kind of logical inference which we use and which is, as the so-called "paradoxes of implication" (among many other problems) show, quite different in some respects from the logical inference conceptualised and formalised by logicians. At the very least, we use a kind of "relevant implication" of the type which Anderson and Belnap have been struggling to capture for the last 15 years; in addition we use rules of modal and deontic inference.

- the nature of defeasibile inference, both in the case of practical argumentation and in general. Defeasible inference is used in many types of argumentation, not only in practical reasoning, but also in moral, deontic and legal reasoning (think of the logic of excuses, for instance). A defeasible inference of the form "if X then(pf) Y" can be invalidated by the consideration of additional relevant facts, and the criteria of relevance vary with the kind of argumentation being pursued and the content of X and Y. A general study of defeasible inference and a typology relevant considerations would be most helpful.

- the nature and rules of our intuitive notions of probability

- the semantics and logic of "reasons" and how they might differ from causes; *

and the list is far from complete!

---

*For an introduction to this area: [Donald Gustafson, "A Critical Survey of the Reasons vs. Causes Arguments in Recent Philsophy of Action", Metaphilosophy, No 4, Vol 4, October 1973].

III. SPEECH-ACT ARGUMENT

At several occasions during the dialogue we are studying, the TD claims that AGEC has not cooperated with the TD in the past, and will probably not do so in the future. The representatives of AGEC first challenge the statement that AGEC has not cooperated in the past (lines 110-113). The TD responds vigorously to this challenge, pointing out that AGEC hired too many lawyers, some of whom tried to apply underhanded political pressure when they could have obtained fair treatment from the TD by approaching it through regular channels - both actions indicative of non-cooperation. (cf. lines 115-143) The AGEC representatives then take the tack of assuring the TD that AGEC does presently want to cooperate with the Government (lines 167-161), and will do so in the future, correcting any failure to cooperate immediately (lines 362-367). The TD does not find these assurances to be convincing, and proposes to put AGEC's avowed cooperativeness to the test, by presenting AGEC with a list of pieces of needed information which AGEC has so far refused to provide (lines 236- 238, 355-360).

The argument which we will be concerned with in this section is the argument the TD seems to use to cast doubt upon the forcefulness of the promise which the AGEC representatives make that AGEC will cooperate with the Government in the future. We call it a "speech-act argument", because it involves querying the conditions for a promise or statement of intent.

i) The Debate Proposition. The debate proposition is that *AGEC wants to cooperate with the Government and will do so in the future.* The AGEC representatives want the TD to adhere to this proposition, but the TD does not. The TD continues to have grave doubts about the debate proposition even after puts forward the argument we will study; this is evidenced by the statement in lines 300-301 and ironical statement in lines 366-367. The following passages are relevant to the establishment of the debate proposition:

******

97 W: As I understand it, the Treasury has taken the position
98 the position that is has simply because it believes that
99 will get the same treatment in the future as it has in the
100 past, in the matter of cooperation from the Associated, in
101 getting information that is necessary on which to compute
102 the tax and then collect it.
103
104 W: And by that you mean we have got no cooperation.
105
106 B: Is that right?
107
108 OLIPHANT: In substance.
109
110 B: If that is correct, it certainly is not in line with my
111 understanding and not in line with the efforts of the
112 Company. We have certainly tried to have our
113 representatives give the Treasury every bit of information.

****

167 HM: Mr dear Mr. Burroughs, may I say this: when
168 Associated Gas and Electric wants to really cooperate with
169 the Government we would like to know it.
170
171 B: You know it right now.

****

193 HM: Mr. Wideman is in charge of this case for the
194 Government, but if the Associated wants to really show that
195 they are cooperating, we would be so pleasantly surprised we
196 would fall over backwards.

****

300 W: I can't go along with the idea that you will cooperate
301 with the Government and are ready and able to pay the tax
302 when due.

****

362 B: And I will say to you, right now, that I will
363 immediately use every effort to see that any failure to
364 cooperate is corrected immediately.
365
366 HM: It will be a very interesting innovation for Associated
367 Gas and Electric.

ii) The Argumentative Utterances. The passage in the dialogue which involves the argument we will study is the following:

****

167 HM: Mr dear Mr. Burroughs, may I say this: when
168 Associated Gas and Electric wants to really cooperate with
169 the Government we would like to know it.
170
171 B: You know it right now.
172
173 HM: And may I ask your position?
174
175 B: I am Vice President
176
177 HM: Are you Attorney?
178
179 B: No, I am not Attorney.
180

181 HM: Are you operating?
182
183 B: No, I am financial officer.
184
185 HM: And Dr. Starch?
186
187 B: Is a director.
188
189 HM: What is his position?
190
191 B: No official position; a director of the company.
192
193 HM: Mr. Wideman is in charge of this case for the
194 Government, but if the Associated wants to really show that
195 they are cooperating, we would be so pleasantly surprised we
196 would fall over backwards.

iii) The Argument. The nature of the argument is determined by lines 167-171. These lines involve a number of relatively complex occurrences, to the analysis of which we now turn.

First of all, Morgenthau says: "...when Associated Gas and Electric wants to really cooperate with the Government we would like to know it.". This might be phrased a bit more explicitly as:

(1) At some future time t1, if AGEC wants to cooperate with the Government at t1, then the TD wants (TD knows (AGEC wants to cooperate with the Government at t1)) at t1.

(One of the contextual pragmatic implications of this is that at the time of the utterance the TD believes that AGEC does not want to cooperate with the Government.)

In the context of the dialogue and its parameters, this statement by the TD is, among other things, a request to AGEC for AGEC to tell the TD when it wants to cooperate.

Burroughs then says in reply: "You know it right now." This might be phrased as follows:

(2) At time t1, TD knows (AGEC wants to cooperate with the Government at time t1).

The presupposition of this is that

(2psp) (AGEC wants to cooperate with the Government at time t1)

and by saying (2), Burroughs has said (2psp), in the weak sense in which one "says" that Q if one says that P, and P presupposes Q.

Now if someone, A, wants to cooperate with someone else, B, at time t, then(pf), A will cooperate with B at time t. (I say "then(pf)" because there might be other

considerations which would lead A not to cooperate, such as over-riding moral rules, or there might be obstacles to cooperation which would prevent A from cooperating successfully.) So if AGEC wants to cooperate with the Government now, then, ceteris paribus, it cooperates with the Government now. And we intuitively feel that present cooperation makes future cooperation more likely. I see two ways of spelling out that intuition, neither of which is really satisfactory, unfortunately. We could make either or both of the following inferences. Since cooperation, especially in the case of legal proceedings, is often an action which takes a relatively long time to perform, it is likely that if AGEC wants to cooperate with the Government at time t1, then it will cooperate with the Government at time t+n (for some limited n). Or: "A cooperates with B" means that A has a relatively enduring positive dispositional attitude towards helping B attain some of B's goals; so if AGEC cooperates with the Government at t, it is likely to cooperate with the Government at t+n (for some limited n).

So if Burroughs were identical with AGEC, then his saying that AGEC wants to cooperate with the Government would be direct testimonial privileged evidence for AGEC's wanting to cooperate. And AGEC's wanting to cooperate would sanction the conclusion that AGEC will cooperate with the Government in the future, by the reasoning sketched in the paragraph above.

Furthermore, if Burroughs were identical with AGEC, then his saying that AGEC wants to cooperate with the Government could count, in context, as *an indirect promise* by AGEC that AGEC will cooperate. The conditions under which a declaration of desire or intent constitute an indirect promise include the conditions under which an utterance of the form "I promise that..." constitutes a promise; they are complex and obscure, but we will simply assume that they would be met in this case. And if someone promises something, then(pf) that will occur.

So if Burroughs were identical with AGEC, then two lines of reasoning would lead to the conclusion that AGEC wants to and will cooperate with the Government.

But, of course, *Burroughs is not identical with AGEC; rather, he is a representative of that corporate body.* One can only have full confidence in statements made by a representative of a corporate body about the desires of that body, if the representative is an *official and fully authorized* representative. And one can only conclude that promises made by a representative are promises made by the corporate body which the representative is representing, if the representative is, again, an official and fully authorized one.

Consequently, the TD, upon hearing Burroughs say "You know it right now", asks Burroughs about his position in the company. It turns out that Burroughs is a Vice-President, but neither an attorney for the company nor an operating officer. In the eyes of the TD, at least, that position is not such as to make him an official and fully authorized representative.

The TD therefore concludes that there is no evidence for the debate proposition, and continues to believe that AGEC does not want to, and will not, cooperate.

Strictly speaking, the argument of the TD's which we will specify does not justify the conclusion that AGEC will not cooperate. Rather, it justifies the conclusion that there is no evidence for thinking that they will cooperate. Since the TD believes at the outset